

SEARCH REQUEST FORM

Access DB# 44400

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Publik
 Searcher Phone #: 308-4740
 Searcher Location: 1ED7
 Date Searcher Picked Up: 6-19
 Date Completed: 6-20
 Searcher Prep & Review Time: 15
 Clerical Prep Time: _____
 Online Time: 15

Type of Search

NA Sequence (#) 10
 AA Sequence (#) _____
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems 05
 WWW/Internet _____
 Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sv model

Run on: June 19, 2001, 17:00:40 ; Search time 7917.7 Seconds
(Without alignments)
1550.862 Million cell updates/sec

Title: US-09-394-519-105
Perfect score: 1299
Sequence: 1 tcttgcacaaataaaaaa.....cagttgttagtaccctaaa 12991

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*
48: gb_est48:*
49: gb_est49:*
50: gb_est50:*
51: gb_est51:*
52: gb_est52:*
53: gb_est53:*
54: gb_est54:*
55: gb_est55:*
56: gb_est56:*
57: gb_est57:*
58: gb_est58:*
59: gb_est59:*
60: gb_est60:*
61: gb_est61:*
62: gb_est62:*
63: gb_est63:*
64: gb_est64:*
65: gb_est65:*
66: gb_est66:*
67: gb_est67:*
68: gb_est68:*
69: gb_est69:*
70: gb_est70:*
71: gb_est71:*
72: gb_est72:*
73: gb_est73:*
74: gb_est74:*
75: gb_est75:*
76: gb_est76:*
77: gb_est77:*
78: gb_est78:*
79: gb_est79:*
80: gb_est80:*
81: gb_est81:*
82: gb_est82:*
83: gb_est83:*
84: gb_est84:*
85: gb_est85:*
86: gb_est86:*
87: gb_est87:*
88: gb_est88:*
89: gb_est89:*
90: gb_est90:*
91: gb_est91:*
92: gb_est92:*
93: gb_est93:*
94: gb_est94:*
95: gb_est95:*
96: gb_est96:*
97: gb_est97:*
98: gb_est98:*
99: gb_est99:*
100: gb_est100:*
101: gb_est101:*
102: gb_est102:*
103: gb_est103:*
104: gb_est104:*
105: gb_est105:*
106: gb_est106:*
107: gb_est107:*
108: gb_est108:*
109: gb_est109:*
110: gb_est110:*
111: gb_est111:*
112: gb_est112:*
113: gb_est113:*
114: gb_est114:*
115: gb_est115:*
116: gb_est116:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est148:*
118: gb_est149:*
119: gb_est150:*
120: gb_est151:*
121: gb_est152:*
122: gb_est153:*
123: gb_est154:*
124: gb_est155:*
125: gb_est156:*
126: gb_est157:*
127: gb_est158:*
128: gb_est159:*
129: gb_est160:*
130: gb_est161:*
131: gb_est162:*
132: gb_est163:*
133: gb_est164:*
134: gb_est165:*
135: gb_est166:*
136: gb_est167:*
137: gb_est168:*
138: gb_est169:*
139: gb_est170:*
140: gb_est171:*
141: gb_est172:*
142: gb_est173:*
143: gb_est174:*
144: gb_est175:*
145: gb_est176:*
146: gb_est177:*
147: gb_est178:*
148: gb_est179:*
149: gb_est180:*
150: gb_est181:*
151: gb_est182:*
152: gb_est183:*
153: gb_est184:*
154: gb_est185:*
155: gb_est186:*
156: gb_est187:*
157: gb_est188:*
158: gb_est189:*
159: gb_est190:*
160: gb_est191:*
161: gb_est192:*
162: gb_est193:*
163: gb_est194:*
164: gb_est195:*
165: gb_est196:*
166: gb_est197:*
167: gb_est198:*
168: gb_est199:*
169: gb_est200:*
170: gb_est201:*
171: gb_est202:*
172: gb_est203:*
173: gb_est204:*
174: gb_est205:*
175: gb_est206:*
176: gb_est207:*
177: gb_est208:*
178: gb_est209:*
179: gb_est210:*
180: gb_est211:*
181: gb_est212:*
182: gb_est213:*
183: gb_est214:*
184: gb_est215:*
185: gb_est216:*
186: gb_est217:*
187: gb_est218:*
188: gb_est219:*
189: gb_est220:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	440	33.9	493	236	AO963001	AO963001 LERGI80TF
2	338.8	26.1	363	188	T14116	T14116 2281 Lambda
3	166.6	12.8	235	236	AO963002	AO963002 LERGI80TR
4	154.4	11.9	540	31	AV527594	AV527594 EST27594
5	142.6	11.0	526	113	AM233999	AM233999 s132h11.Y
6	142.6	11.0	558	118	AM597702	AM597702 s149g05.Y
7	136.2	10.5	448	118	AM620835	AM620835 s147b12.Y
8	135.8	10.5	534	137	BE611146	BE611146 sq/6d03.Y
9	134.6	10.4	587	113	AM257000	AM257000 EST305137
10	134.6	10.4	791	155	BC586052	BC586052 EST487817
11	127.4	9.8	485	238	AE127897	AE127897 OSJNB009
12	126.2	9.7	270	31	AV549228	AV549228 AV549228
13	120.6	9.3	532	16	A1161927	A1161927 A009P620
14	119	9.2	571	103	A1894846	A1894846 EST264289
15	119	9.2	766	174	BC130539	BC130539 EST476185
16	118.8	9.1	385	118	AM622377	AM622377 EST313176
17	118.4	9.1	692	155	BC592323	BC592323 EST491001
18	117	9.0	522	21	A1490285	A1490285 EST248611
19	116.2	8.9	611	111	AM065881	AM065881 687002C04
20	115.4	8.9	574	31	AV550536	AV550536 AV550536
21	114.8	8.8	327	153	BC405537	BC405537 sac35e02
22	113.8	8.8	475	144	BF097661	BF097661 EST415734
23	113.8	8.8	583	21	A1487752	A1487752 EST246074
24	112.6	8.7	796	151	BF622756	BF622756 HVSME000
25	112.4	8.7	512	118	AM585659	AM585659 EST317282
26	111.8	8.6	658	119	AM696439	AM696439 NF106B07
27	111.4	8.6	373	153	BC405482	BC405482 sac44a11
28	110.4	8.5	348	17	A1165295	A1165295 A080P54U
29	107.2	8.3	663	105	AL507122	AL507122 AL507122
30	107	8.2	429	30	AV409036	AV409036 AV409036
31	106.6	8.2	559	158	H77029	H77029 17460 Lambda
32	106	8.2	723	155	BC586298	BC586298 EST488064
33	105.4	8.1	522	137	BE555532	BE555532 sp90e07.Y
34	102.6	7.9	564	143	AM720051	AM720051 LJNEST3A
35	101.6	7.8	434	149	BF010257	BF010257 sas9e07.Y
36	100.6	7.7	489	110	AM031324	AM031324 EST274778
37	100.4	7.7	590	20	A1461184	A1461184 sa76b10.Y
38	100.2	7.7	420	153	BC405625	BC405625 sac45e11
39	100.2	7.7	479	145	BF203107	BF203107 WHE1759-1
40	100.2	7.7	480	145	BF202003	BF202003 WHE1759-1
41	100	7.7	499	237	AZ044889	AZ044889 A363/F.SO
42	100	7.7	514	234	AO842107	AO842107 T134402.S
43	100	7.7	588	234	AO842067	AO842067 T134353.S
44	99.6	7.6	554	136	BE500188	BE500188 WHE0980.F
45	99	7.6	404	136	BE474893	BE474893 sp69g07.Y

ALIGNMENTS

RESULT	1	493 bp	DNA	GSS	28-JAN-2000
LOCUS	AO963001/c				
DEFINITION	LERGI80TF LERG Arabidopsis thaliana genomic clone LERG180, DNA				
ACCESSION	AO963001				
VERSION	AO963001.1				
KEYWORDS	GI:6790702				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Arabidopsis thaliana				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis; 1 (bases 1 to 493)				
TITLE	Buell, C.R., Lin, X., Pal, G., Barnstead, M., Bowman, C., Utechtach, T., Feldbljum, T., Liang, F., Creasy, T., and Fraser, C.M.				
	Genomic survey sequencing of Landsberg erecta ecotype of				

JOURNAL COMMENT

Arabidopsis thaliana and identification of sequence-based polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: atel@igf.org
For additional information, see <http://www.tigr.org/cdb/at.html>
Seq primer: TF
Class: shoigen

FEATURES

Source
Location/Qualifiers
1..493
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERGI80"
/clone_lib="LERG"
/note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."

BASE COUNT

179 a 100 c 79 g 135 t

ORIGIN

Query Match 33.9%; Score 440; DB 236; Length 493;
Best Local Similarity 100.0%; Pred. No. 3.6e-108;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	858	ccatttgagagcctcctcaactccagttccatccatcaagcaccatagttggaat	917
DB	492	ccatttgagagcctcctcaactccagttccatccatcaagcaccatagttggaat	433
QY	918	ggcaatcgaagaagaataatgatgatgatgatgatgatgatgatgatgatgat	977
DB	432	ggcaatcgaagaagaataatgatgatgatgatgatgatgatgatgatgatgat	373
QY	978	gaagctaccacatcgttgggttgggttgggttgggttgggttgggttgggttgc	1037
DB	372	gaagctaccacatcgttgggttgggttgggttgggttgggttgggttgggttgc	313
QY	1038	gcattcgttggttcattgacataatattttagtcattgacacccatccatgagtt	1097
DB	312	gcattcgttggttcattgacataatattttagtcattgacacccatccatgagtt	253
QY	1098	aagaaatgattacatttttaccatgtttttaaagaatgagaggttaactataatc	1157
DB	252	aagaaatgattacatttttaccatgtttttaaagaatgagaggttaactataatc	193
QY	1158	tgggtgcttacctattttagctcttcttcttcttcttcttcttcttcttcttct	1217
DB	192	tgggtgcttacctattttagctcttcttcttcttcttcttcttcttcttcttct	133
QY	1218	attaacttggttggttggttggttggttggttggttggttggttggttggttg	1277
DB	132	attaacttggttggttggttggttggttggttggttggttggttggttggttg	73
QY	1278	ctcagttggttggttggttggttggttggttggttggttggttggttggttg	1297
DB	72	ctcagttggttggttggttggttggttggttggttggttggttggttggttg	53

RESULT	2	363 bp	MRNA	EST	07-JAN-1998
LOCUS	T14116				
DEFINITION	2281 Lambda PR2 Arabidopsis thaliana cDNA clone 47E1077, mRNA				
ACCESSION	T14116				
VERSION	T14116.1				
KEYWORDS	GI:931090				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 363)

REFERENCE
AUTHORS
Newman, T., deBruin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Ralke, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.

TITLE
JOURNAL
MEDLINE
COMMENT
95148729

On Nov 29, 1993 this sequence version replaced gi:430485.

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRC, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 223j3tcn@lhm.cl.msu.edu
Seq primer: 77.

FEATURES
source
1. 363
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="47E1077"
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: NotI; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT
106 a 90 c 73 g 85 t 9 others

ORIGIN

Query Match 26.1%; Score 338.8; DB 188; Length 363;
Best Local Similarity 96.7%; Pred. No. 8e-81;
Matches 351; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1 tcttgcacaaataaataaagaaggttggtagaagcaatgaacgaatcgtga 60
DB 1 TTTTCCCAAAATTAATAAATAAAGAGTGTGTGAAGCCAAATGAACGAGATGCTGA 60
QY 61 caaacacatgctaaacacacagtcgaatcctctcctgctactcctgtgaaagaa 120
DB 61 CAAACACATGCGCTAAACAAACAGTCGATCCTCTCTCTGCTACTCTGCGAAGGA 120
QY 121 aggcgaacccggaagaagcaggttgaactgtccaagatgcaactccttaacacca 180
DB 121 AGGCAAGACCGGGAAGACCAAGGCTTTGACATGTCACAGATGCAACTCTTAAACACA 180
QY 181 agttcgttactaacaactacagcctgaagccagcaggttacttttgaagactgca 240
DB 181 AGTTGTTACTACAACTACAGCTGACGAGCCAGGTACTTTGTAAGACTGCA 240
QY 241 ggaagatggaacgaggtgttccctcagaagacacccgcgtcgggtgagcgagtcgca 300
DB 241 GGAGGATTTNGACCGCAGTGGGTCTCCCTCAGGACATCCCGCGGTGGCGAGTCCGA 300
QY 301 agaacagaagatcttcccaatcctcctcctcctacacacttc-gtcttctctca 359
DB 301 AGAACAGAAGATCTTNTTCCATTCCTCTTCTGACCCCTTNGTCTTCTTCTTNA 360
QY 360 aag 362
DB 361 AAG 363

RESULT 3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AO963002 255 bp DNA GSS 28-JAN-2000
LEK8180TR LERG Arabidopsis thaliana genomic clone LERG180, DNA sequence.
AO963002
AO963002.1 GI:6790703
GSS

thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 255)

Buell, C.R., Lin, X., Pal, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblyum, T., Liang, F., Greasy, T. and Fraser, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ateflgr.org
For additional information, see <http://www.tigr.org/tdb/ac/at.html>
Seq primer: TR
Class: shotgun.

FEATURES
source
1. 255
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERG180"
/note="Organ: Leaf; Vector: pUC19UK; Total genomic DNA was sheared to 0.4-0.7 kbp before ligation."

BASE COUNT
60 a 61 c 58 g 75 t 1 others

ORIGIN

Query Match 12.8%; Score 166.6; DB 236; Length 255;
Best Local Similarity 90.7%; Pred. No. 2.8e-34;
Matches 223; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 658 ctctccttaactcaagccctgcttcaatgtctatccaaatgttcaatgattcattgct 717
DB 1 CTTCCTTAACACACCCCTGCTTCAATGCTTATCAATGTCATGATTCATCGCTCT 60
QY 718 aacacacatcgaggttggctgaagttaccacagtttaagatcagagacagctt 777
DB 61 ACACACATCGGGGTTGGTGTGAGTTACCCACAGTTTAAAGATTCAAGACACAGCTT 120
QY 778 tggagttctctctgaagtgaggtatccctacgtcaagaagaggggtccagtgacata 837
DB 121 TGGGATTCCTCTGTATGTTGGGATCCCTA-NTCAAAAGAGGGGTCCAGT-GCAGTA 178
QY 838 ataattggaagccggtgtgcatcttgagaagcctcctcaaaacttccagttcatcatca 897
DB 179 ATAAATGAGGCC---GTGCTGATTTGAGAG---CCTTAAACTTCAGTTTCATCATCA 232
QY 898 gcaacca 903
DB 233 GCCCAA 238

RESULT 4
LOCUS
DEFINITION
AV527594 540 bp mRNA EST 01-SEP-2000
AV527594 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP240f09R 5', mRNA sequence.

ACCESSION	AV527594
VERSION	AV527594.1
KEYWORDS	GI:8687122
SOURCE	EST.
ORGANISM	Thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1 (bases 1 to 540) Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S. A large scale analysis of cDNA in Arabidopsis.thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research, Kazusa DNA Research Institute Yama 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamiz@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/ Location/Qualifiers
FEATURES	1..540 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="AP240F09R" /clone.lib="Arabidopsis thaliana aboveground organs two to six-week old" /tissue.type="aboveground organs"; /dev_stage="two to six-week old" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	175 a 124 c 110 g 131 t
ORIGIN	
Query Match	11.9%; Score 154.4; DB 31; Length 540;
Best Local Similarity	67.0%; Pred. No. 6.9e-31;
Matches 240; Conservative	0; Mismatches 106; Indels 12; Gaps 1;
Oy	23 aaagaagttgttgggaagccaatgaagaagtgcgtgcaaacacatgccttaacaaca 82
Dd	151 ACAGGAGATTGTAGTGAAGGCCCTTGGAAGAATACTAACAAACACATGCCCAAGCCGA 210
Oy	83 gtccgatcctccttccttcctgcctacc-----tgtgaaagaaggaagaagcc 130
Dd	211 ACCGCACCGGTTCAACGGCAGCACGCCACCOCCTGGGTGGGTGAGAGAGAACGCAAGGCC 270
Oy	131 ggagaagaaccaaggctttggaactgtgccaaagtatgcacctctaaccaccaagtttgta 190
Dd	271 AGAAAAAGATCAAGCTGTAAACTGTGCCAGATGTAACCAACCAACAAGCTTTGTGA 330
Oy	191 ctacacacactacaagcccgacgcgaagcccgagtaactttgtaaagactcagaaggtattg 250
Dd	331 CTACAAACAATTATATGTTGAGCGCACCCAAGATACCTTCGCAAAAGTTGTGAGAAGTATTG 390
Oy	251 gaccgcagtggttcccctcagaagaacatccccogtcgtygagcgycgtccgcaagaagaag 310
Dd	391 GACCCAAAGCGGTTGGCTTAGAAMCATTCCTGTTGGCGGTGGCTCTAATAAACAACAAGAG 450
Oy	311 atctcttcacatctcctctcctctcaccctcctcctcctcctctcttcttaagaagaacct 368
Dd	451 ATCTACACTTCTTCTTCTGTGAATTAGTAACAACTACTCGGATTTCTACACAACGACT 508
RESULT	5
LOCUS	AW233999 526 bp mRNA EST 17-JUL-2000
DEFINITION	sfs32h11.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl028-1726 5' similar to TR:082027 082027 DOF ZINC FINGER PROTEIN ;, mRNA sequence.
ACCESSION	AW233999
VERSION	AW233999.1 GI:6566326

KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	Glycine max
REFERENCE	Eurytota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
AUTHORS	1 (bases 1 to 526) Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
JOURNAL	Public Soybean EST Project
COMMENT	Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Insert Length: 1210 Std Error: 0.00 High quality sequence stop: 405. Location/Qualifiers 1..526 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1726" /clone_lib="Gm-cl028" /tissue_type="roots of 'Supernod' plants" /lab_host="DH10B" /note="Vector: plbuescript II XR; Site.1: EcoRI; Site.2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hememethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A,C, or G) was added to the 3' end of the primer (GAGAGGAGAGAGAGAGAACTAGCTGAG(7)18V) to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI: all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA size fractionation column. The column eluent was then ligated into Stratagene's plbuescript II XR predigested vector (plbuescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Corvelli."
BASE COUNT	144 a 146 c 102 g 134 t
ORIGIN	
Query Match	11.0%; Score 142.6; DB 113; Length 526;
Best Local Similarity	71.6%; Pred. No. 11e-27;
Matches 187: Conservative	0; Mismatches 74; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:51 ; Search time 7917.7 Seconds
(without alignments)
1457.738 Million cell updates/sec

Title: US-09-394-519-127

Perfect score: 1221
Sequence: 1 atttccttcacaaagagtl.....aaaaaaaaaaaaaaaaaa 1221

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: gb_est41.*
42: gb_est42.*
43: gb_est43.*
44: gb_est44.*
45: gb_est45.*
46: gb_est46.*
47: gb_est47.*
48: gb_est48.*
49: gb_est49.*
50: gb_est50.*
51: gb_est51.*
52: gb_est52.*
53: gb_est53.*
54: gb_est54.*
55: gb_est55.*
56: gb_est56.*
57: gb_est57.*
58: gb_est58.*
59: gb_est59.*
60: gb_est60.*
61: gb_est61.*
62: gb_est62.*
63: gb_est63.*
64: gb_est64.*
65: gb_est65.*
66: gb_est66.*
67: gb_est67.*
68: gb_est68.*
69: gb_est69.*
70: gb_est70.*
71: gb_est71.*
72: gb_est72.*
73: gb_est73.*
74: gb_est74.*
75: gb_est75.*
76: gb_est76.*
77: gb_est77.*
78: gb_est78.*
79: gb_est79.*
80: gb_est80.*
81: gb_est81.*
82: gb_est82.*
83: gb_est83.*
84: gb_est84.*
85: gb_est85.*
86: gb_est86.*
87: gb_est87.*
88: gb_est88.*
89: gb_est89.*
90: gb_est90.*
91: gb_est91.*
92: gb_est92.*
93: gb_est93.*
94: gb_est94.*
95: gb_est95.*
96: gb_est96.*
97: gb_est97.*
98: gb_est98.*
99: gb_est99.*
100: gb_est100.*
101: gb_est101.*
102: gb_est102.*
103: gb_est103.*
104: gb_est104.*
105: gb_est105.*
106: gb_est106.*
107: gb_est107.*
108: gb_est108.*
109: gb_est109.*
110: gb_est110.*
111: gb_est111.*
112: gb_est112.*
113: gb_est113.*
114: gb_est114.*
115: gb_est115.*
116: gb_est116.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SOURCE	ORGANISM	thale cress.
REFERENCE	Arabidopsis thaliana	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 443)	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries	
JOURNAL	DNA Res. 7, 175-180 (2000)	
MEDLINE	20363093	
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. Location/Qualifiers	
FEATURES	Source	1..443
	/organism="Arabidopsis thaliana"	
	/strain="Columbia"	
	/db_xref="taxon:3702"	
	/clone="SQ1721a1f"	
	/clone_lib="Arabidopsis thaliana green siliques Columbia"	
	/tissue_type="green siliques"	
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	136 a 107 c 89 g 111 t	
ORIGIN		
Query Match	31.9%; Score 389; DB 31; Length 443;	
Best Local Similarity	94.1%; Pred. No. 8.1e-87;	
Matches 417; Conservative	0; Mismatches 20; Indels 6; Gaps 1;	
OY	732 tgcgaatacttgaacacggtgctgtagaagaagctcaaggaagaacacggaagcttcaaaa	791
Db	443 TCGGAATCTTGAACGCGTCTGCTAGAGAACCTACGGAAGAGAACCGAGCTTGAANA	384
OY	792 gagcgtatgagcttcgaactcgaactcgaactcgaactcgaactcgaactcgaactcga	851
Db	383 GAGCGTATGAGAGCTTCGAACCTCAAGCGTCTCCACAACTCTACGGTCAGATGACTCA	324
OY	852 ccaactaacatcatatgtgtccttcctgagagaggtgtagctggtccatcatcagcaac	911
Db	333 CCAACTAACCTCATATGTCGTCTGTCGCGAGCGTGGGCGCCATCATCATCGAAGC	264
OY	912 catcccaacaatcacaagccggttccgattcaaccgctgagatgtgtgtgtcgtcaggt	971
Db	263 CATCAACCAATACACAGGCCGCTTCTATCAATCCGTGGGTCCTGTGCTGTCAGGTC	204
OY	972 gtcctatgagctgaacttttgaagccttcgctcagatcgtaat-----tttaagtgt	1025
Db	203 GCTCATGAGCGCTCAATTTTGAACCTTGGCTCCACATCGTGAATTTTATTTTAACTGG	144
OY	1026 gagggaaggtgtgttgggttttttcaatcatcgttatagctatcgtgtgtgggtcatt	1085
Db	143 GGAAGAGGCTTTTGGTATTTTTCGTTATCTTATATAGTCTATCTGTGGGCTCATTT	84
OY	1086 gtaattttgagatgttgcccttcatactgaactagtcatactgatatgatatgaactta	1145
Db	83 GTAATTTTGGATGATGGCCCTTCATCATGACATGATGATGATGATGATGATGATGATG	24
OY	1146 tatttcaagtgcacaaacttaatt 1168	
Db	23 GATTTAAATTTAGCAAAAATTAAGT 1	
RESULT	3	
LOCUS	T43527	538 bp mRNA EST 07-JAN-1998
DEFINITION	6790 Lambda-PRL2 Arabidopsis thaliana cDNA clone 120N14T7, mRNA	
	sequence.	

ACCESSION	T43527
VERSION	T43527.1
KEYWORDS	GI-2758376
SOURCE	EST.
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons: core eudicots: Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi 1 (bases 1 to 358) Newman,T., deRuijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh ,L., Ollivierre,J., Ralkehl,N., Somerville,S., Thomashow,M., Retzel ,E. and Somerville,C.
AUTHORS	
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
JOURNAL	95148729
MEDLINE	On Jan 7, 1998 this sequence version replaced g1:947926.
COMMENT	Contact: Thomas Newman Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel.: 517-353-0854 Fax: 517-353-9168 Email: 22313cne@dm.cl.msu.edu Seq primer: 77.
FEATURES	Location/Qualifiers
source	1..538 /organism="Arabidopsis thaliana" /strain="var columbia" /db_xref="taxon:3702" /clone="12ON1477" /clone_1ib="Lambda-PRL2" /note="vector: lambda ZAP-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark - rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BR's lambda ZAP-lox. The cDNA inserts were directionally cloned with Sal-I Not arms using oligo dt primed cDNA."
BASE COUNT	143 a 114 c 129 g 131 t 21 others
ORIGIN	
Query Match	31.2% Score 380.4; DB 189; Length 538;
Best Local Similarity	93.3%; Pred No. 1.2e-84;
Matches 434; Conservative	0; Mismatches 27; Indels 4; Gaps 4
Oy	1 atttcttcacaaagagctcctaactcgagtgaagaacaaccatttcataccta 60 Dbd 1 ATTTCCTCCCAACAAGACTCCTAAGTGAAGAACAACACATTCTNATCTCA 60 Oy 61 tctcgaagaacaacaaccttcgtgltcttccttcctatctcataaggaatat 120 Dbd 61 TCTCGAAAGAACCAACCATTTGCGTGCTCTTCTNTCTCTATTCTCATTAAGCAATAATA 120 Oy 121 ttccctgaacgttgttgcttctgtgaagaataaaaaaatgatgatgggcaagaag 180 Dbd 121 TTCCTGAACACTTGTGATTCGTGTGAAAAGAAATAAAAACATGATGATGGCGCAAGAG 180 Oy 181 acctgagtttagcctaagcttagggttttaacaanaatacacatcctctcagaataac 240 Dbd 181 ATCTTAGGTTTGACCCTAACCTTAGGGCTTTTCAACAAATACATCAANTCTCTTCCAGATGAATC 240 Oy 241 tgaattcctaactcttcatacatcaaaaacatctccagagactcccatggaaccacaacttgc 300 Dbd 241 TGAATTCCTAAGCTCTCATATATCAACAACATCTCCAGAGACTCCCATSGAACAACATTCGC 300 Oy 301 atccatcatcagatcttcgacaagatatagac-gtgaacagtttccatcaaac-gttaactg 358

Db 301 NTCCTACATGATCTTGGCAAGNTAGCGGTGACCACTTTTCATCAACGGTTAACTG 360

Qy 359 CGAGGAAGACACAGAGAGTTCCTGTCACCAAGTACGATCAGCAGCAGGAGG 418

Db 361 CGAGNAGGACAGAGAGTTCCTGTCACCAAGTACGATCAGCAGCAGGAGG 419

Qy 419 GAGAGTGAAGAGAGAGAGTTCCTGTCACCAAGTACGATCAGCAGCAGGAGG 463

Db 420 GGGAGGTTGNNAG-AGCGATTTCCGACCCGCGTGGTCCGG 463

RESULT 4

Al100243 545 bp mRNA EST 21-AUG-1998

LOCUS 34618 Lambda-PRL2 Arabidopsis thaliana cDNA clone 110C16XP 3', mRNA

DEFINITION sequence.

ACCESSION Al100243

VERSION Al100243.1 GI:3450204

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 545)

AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.

TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones

JOURNAL Plant Physiol. 106, 1241-1255 (1994)

MEDLINE 95148729

COMMENT Contact: Thomas Newman

MSU-DOE Plant Research Laboratory

Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI

Tel: 517-353-0854

Fax: 517-353-9168

Email: 22313tcm@cl.msu.edu

The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation.

Seq primer: M13_Universal.

FEATURES

source

1. 545

/organism="Arabidopsis thaliana"

/strain="var columbia"

/db_xref="taxon:3702"

/clone="110C16XP"

/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is PRL's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

BASE COUNT 129 a 116 c 131 g 163 t 6 others

ORIGIN

Query Match 29.88; Score 364; DB 16; Length 545;

Best Local Similarity 90.88; Pred. NO. 1.5e-80;

Matches 423; Conservative 0; Mismatches 32; Indels 11; Gaps 3;

Qy 712 TAAAGCAACGAGGATGCTGCAATCTGCAAGAGTGTGAGAGAGTAAAGAG 771

Db 85 TTAAGNAAACGAGGTAGATT---CGATACTGAACGCTGCTAGAGAGTAAAGAG 140

Qy 772 AGAACCGAGAGCTCAGAGAGGTATGAGCTTCAACTCAAGCTGTCTCACAAT 831

Db 141 AGACCCGAGAGCTCCAG-AGAGGCTATGAGCTTGCAGACTCAAGCTGTCTCACAAT 199

Qy 832 TCTAGTCAATGATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 891

Db 200 TCTAGGCTCAATGATCACTCACTCACTCACTCACTCACTCACTCACTCACT 259

Qy 892 CTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 951

Db 260 GTGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 319

Qy 952 TTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1011

Db 320 TTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 379

Qy 1012 aat-----ttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1065

Db 380 GATTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 439

Qy 1066 TCTATCTGT 1125

Db 440 TCTATCTGT 499

Qy 1126 GTATGATGCAACCTTAAAGATTAAATTGACNAAATTAGTTAC 1171

Db 500 GTATGATGCAACCTTAAAGATTAAATTGACNAAATTAGTTAC 545

RESULT 5

T42030 475 bp mRNA EST 07-JAN-1998

LOCUS 5293 Lambda-PRL2 Arabidopsis thaliana cDNA clone 110C167, mRNA

DEFINITION sequence.

ACCESSION T42030

VERSION T42030.1 GI:932917

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 475)

AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.

TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones

JOURNAL Plant Physiol. 106, 1241-1255 (1994)

MEDLINE 95148729

COMMENT On Nov 29, 1993 this sequence version replaced gi:534618.

Contact: Thomas Newman

MSU-DOE Plant Research Laboratory

Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI

Tel: 517-353-0854

Fax: 517-353-9168

Email: 22313tcm@cl.msu.edu

Seq primer: T7.

FEATURES

source

1. 475

/organism="Arabidopsis thaliana"

/strain="var columbia"

/db_xref="taxon:3702"

/clone="110C167"

/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and

siliques. The vector is BRL's lambda Zip-Tox. The cDNA inserts were directionally cloned with Sal-I Not arms using oligo dt primed cDNA. "

BASE COUNT 144 a 99 c 111 g 103 t 18 others

ORIGIN

Query Match 29.8%; Score 363.4; DB 188; Length 475;
Best Local Similarity 88.2%; Pred. No. 2.1e-80;
Matches 419; Conservative 0; Mismatches 49; Indels 7; Gaps 3;

QY 148 aggaataaaacatgatgatggcaagaagatcagtttgagcctaagcttaggt 207
DB 1 AGGAATAATAAAAACATGATGATGGCAAGAGATCTAGGTTGAGCCTTAAGCTTAGGGT 60
QY 208 ttccaaaataacacatcctcttcagatgaatcgaatccttaacttcttaataaac 267
DB 61 TTTCACAAATCCAAATCTCTTCATGATGATGATCTTAATCTTCTTCTTATCAACA 120
QY 268 atcccaagactcccatggaaccaacatcgcctcatalcatalcatalcatalcatalc 327
DB 121 ATCTCCAGAGACTCCCATGGAACCAACATTCGATCTCATCATCATCATCATCATCAT 180
QY 328 acgttgacagtttccatcaacaggttaactgcgaggaagacacaggaagttcgtaccaa 387
DB 181 ACCTGGAACAGTTTTCATCAACGTTTAACTGCGAGAGACACAGAGATTTCGTACCAA 240
QY 388 acagtaacatcaccac 447
DB 241 ACAGTACGATCTCAACACACATTTAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 448 ccgagcttgctcccgagac 505
DB 301 CCGGCGTTGGCTCGGCGAG 360
QY 506 tggagacctaagat----gaagaagaacggggcggaacgctcgaggaagaacacacacac 561
DB 361 TGGAACTCAGATGAG 420
QY 562 tatcaagaatcagctgcttcttcgaagaagacttcaagaacacacacacacacacac 616
DB 421 TTTTCAAGATCGNCTCTTTTCTTTTGG-GGGGTTTTCANNNCAANANNTT 474

RESULT 6
AV556109/c 375 bp mRNA EST 06-SEP-2000
LOCUS AV556109 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION thaliana cDNA clone S003490ZF 3', mRNA sequence.
ACCESSION AV556109
VERSION AV556109.1 GI:8727524
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 375)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..375
/organism="Arabidopsis thaliana"
/strain="Columbia"

/db_xref="taxon:3702"
/clone="S003490ZF"
/clone.lib="Arabidopsis thaliana green siliques Columbia"
/tissue_type="green siliques"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 126 a 84 c 78 g 87 t

ORIGIN

Query Match 26.2%; Score 319.4; DB 31; Length 375;
Best Local Similarity 92.8%; Pred. No. 2e-69; Indels 6; Gaps 1;
Matches 348; Conservative 0; Mismatches 21;

QY 803 gctcgaactcacaagctgtccacacatcaggtcagatgactcccaactacact 862
DB 375 GCTTCAACTCTCAAGCTGTCTCCAAATTCAGGTGATGATGATGATGATGATGATGAT 316
QY 863 catcatgtgctcctcgtcgcgagctgtagctggtccatcatcatcatcatcatcatcat 922
DB 315 CATCATGTCTCTCTCTGCGAGAGCTGTGGTGCCCATCATCATCATCATCATCATCAT 256
QY 923 tcaabagccggtttgattaaaccggtggttctgtgtcaggtggtcagtggt 982
DB 255 TCACAGGCCGTTTCTATCAATCCGTGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 196
QY 983 gaatttgaagccttgctccacacacacacacacacacacacacacacacacacacac 1036
DB 195 GAATTTGAAGCCTTCTGCTCCAGATCGATGATTTTATTTAGTGTGGAAAGAGGCT 136
QY 1037 ttltgggtttttcattatcgtatatactatcctcgtgtgggtgcatgttaatttga 1096
DB 135 TTTTGTATTTTTCATTATGCTATATATATATATATATATATATATATATATATATAT 76
QY 1097 tgaattgcttctcattgaactgcatatgatgatgacacacttaaatatttcaagta 1156
DB 75 TGATTTGGCTTCTCATGAATATGCTATGATGATGATGATGATGATGATGATGATGAT 16
QY 1157 gcaaaacttaattac 1171
DB 15 GCAAAATTTAGTTAC 1

RESULT 7
AI099721 389 bp mRNA EST 21-AUG-1998
LOCUS AI099721
DEFINITION 33874 lambda-PRL2 Arabidopsis thaliana cDNA clone 120N14AP 3', mRNA
sequence.
ACCESSION AI099721
VERSION AI099721.1 GI:3449460
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 389)
Newman, T., debruin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh
, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel
, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PR, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313t@dm.cl.msu.edu
The sequence entry for this EST has been reverse complimented and

JOURNAL
MEDLINE
COMMENT

large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22333tcn@lhm.cl.msu.edu

Seq primer: T7 dye primer.
Location/Qualifiers

FEATURES
source

1. 390

/organism="Arabidopsis thaliana"
/strain="var.columbia"
/db_xref="taxon:3702"
/clone="200B97"

/note="Vector: lambda Zip-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

BASE COUNT 94 a 79 c 95 g 107 t 15 others
ORIGIN

Query Match 23.5%; Score 287; DB 158; Length 390;
Best Local Similarity 86.7%; Pred. No. 2.5e-61;
Matches 333; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 705 accaagttaaagcaacgagtagatctgcgaacttgaaacgctgctagaagaacta 764
|||||
Db 1 ACCAACTTAAGCAACGAGGAGTAGATTCGAATTAAGTGAACGCGTAGAAGCACTA 60
QY 765 agcgaagagacacgagacttcagaaagagctatgagacttcgaactcgaagctgct 824
|||||
Db 61 ACGGAAGANAACCGGAGACTTCAGAAAGAGCGCTAGAGCTTCAGACTCAGCTGCT 120
QY 825 ccacactctacgctgctgactgcacacacactcactcactcactcactcactcactc 884
|||||
Db 121 CCACAACTTNTGCGTAGATGACTCCACCAACTACACTCATCTCATCTTCTGCGAG 180
QY 885 cgtgtagctgctcactcactcactcactcactcactcactcactcactcactcactc 944
|||||
Db 181 CMTNTGGTGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 240
QY 945 ccgtgtagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1003
|||||
Db 241 CCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 1004 acgactgtaattttt-----agtgtagggggaagagtggttttggttttttcattac 1056
|||||
Db 301 ACGNTGGAATTTTNTTNTTAAGTGTGAGGAAAGGAGGTTTGGGATTTTGGGNATC 360
QY 1057 gttatatagctactcgtgctgctgctgctgctgctgctgctgctgctgctgctg 1080
|||||
Db 361 GNAANATGCTTNTCTTGTGGG 384

RESULT 10
LOCUS BG441022 860 bp mRNA EST 15-MAR-2001
DEFINITION GA_Ea0011F23f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0011F23f, mRNA sequence.
ACCESSION BG441022
VERSION BG441022.1 GI:13350674

KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
Rosidae: eudicots II: Malvales: Malvaceae: Gossypium.

REFERENCE 1 (bases 1 to 860)
Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, D., Henry

D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution

of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: TAATGACTCCTATGAGG
High quality sequence stop: 746.

FEATURES
source

1. 860

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0011F23f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/clone_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 257 a 190 c 222 g 186 t 5 others
ORIGIN

Query Match 19.9%; Score 243.4; DB 153; Length 860;
Best Local Similarity 67.3%; Pred. No. 2.4e-50;
Matches 406; Conservative 0; Mismatches 161; Indels 36; Gaps 3;

QY 299 cgatcctacatcagatcttcgcaagatagacgtgcaagcttccatcgaagcttaactg 358
|||||
Db 7 CGAGACTATGATCTTTCTTAAAGAGATTCACGAGCAATTCGATTCATCGTCCATG 66
QY 359 cgaaggaagacacagaggttctgcaccacacagtagatcctaagacattagcgggaa 418
|||||
Db 67 CGAAGAGAGAGCTGGAGTTTCACTTCGACACGACGATATCGAG--TGTGAGTGGAA 123
QY 419 gagaagctgagaggaagaaatcgcggaacgcgctgctccgcgcagatcagcagca 478
|||||
Db 124 AAGGAGCGAAGGGAAGCGCATTAATGGA-----GATCA 156
QY 479 gatactccgagatcgaggttaactacgctggaactatagaaagaagaagcggggcga 538
|||||
Db 157 GCTTGATATTTAAAGAGCTGTTCTGCTGTCATCATGTATGAGAGAT-----GCTGA 210
QY 539 aacgtgaggaagaacactaggtatcacaagaatcagctgcttctcgaagaagactt 598
|||||
Db 211 CGCTTGAGAAAAAACTTAAGCTTTTCGAAGATCAGTCTGCTATTCTTGAAGAAAGCTT 270
QY 599 caaagaacaacaactcactcactcactcactcactcactcactcactcactcactcact 658
|||||
Db 271 TTAAGAACACAACTCTGACCAACCAAAAGAGAGGCTTTGGCTTAACAGCTGGAGAT 330
QY 659 gacgcagacagaagtggaagtgtgtccaaaacagaagagctagaaaccaagttaagca 718
|||||
Db 331 GCGAGCCAGACAGAGTTGAAGTTGGTTCCAAAAACAGAGGCGAGAGCCAGCTGAGCA 390
QY 719 aacgagagtagatgtggaacttgaacggtgctgagagaagcctaagcgaagaagacgc 778
|||||
Db 391 AACTGAGCTTGAAGTCTTAAGAGAGATGCTGTGAGAAATCTGACGAGGAGAAACAG 450

QY 779 gagacttcagaagaagctatgagcttcgaactcgaagctgtctccacaattctacgg 838.
 DB 451 GCGGTTCCAGAGAAAGTTAAGAGCTGAGAGCAGTGAATTTCTCTCCTCAGTTCTACAT 510
 QY 839 tcagatgactccaccactacactcatcatgtctcctcgtgagaggttgaagctgtcc 898
 DB 511 GCAAAATACCCCACTTACGACCTCACCAGTGTGCCCATCGTGTGAGAGAGTAGGTGATC 570
 QY 899 atc 901
 DB 571 ATC 573

RESULT 11
 AM727255 standard; RNA; EST; 868 BP.
 ID AM727255
 AC AM727255;
 SV AM727255.1
 DT 24-APR-2000 (Rel. 63, Created)
 DT 18-NOV-2000 (Rel. 65, Last updated, Version 2)
 DE GA_Ea001f23 Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum
 DE cDNA clone GA_Ea001f23, mRNA sequence.
 XX EST.
 KM Gossypium arboreum
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Malvales; Malvaceae; Gossypium.
 OC [1]
 RP 1-868
 RA Wing R.A., Frisch D., Yu Y., Main D., Rambo T., Simmons J., Henry D.,
 RA Wood T.C., Leslie A., Wilkins T.A.;
 RT "An integrated analysis of the genetics, development, and evolution of the
 RT cotton fiber";
 RL Unpublished.
 XX
 CC Contact: Wing RA
 CC Clemson University Genomics Institute
 CC Clemson University
 CC 100 Jordan Hall, Clemson, SC 29634, USA
 CC Tel: 864 656-7288
 CC Fax: 864 656-4293
 CC Email: rwing@clemson.edu
 CC High quality sequence stop: 868.
 CC
 XX
 FH Key Location/Qualifiers
 FH
 FT source 1. 868
 FT /db_xref="taxon:29729"
 FT /db_xref="ESTLIB:2480"
 FT /note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"
 FT /organism="Gossypium arboreum"
 FT /strain="AKA"
 FT /cultivar="8400"
 FT /clone="GA_Ea001f23"
 FT /tissue="Gossypium arboreum 7-10 dpa fiber library"
 FT /tissue_type="Fibers isolated from bolls harvested 7-10
 FT dpa"
 FT /lab_host="E. coli"
 XX
 SO Sequence 868 BP; 259 A; 192 C; 226 G; 186 T; 5 other;

Query Match 19.94; Score 243.4; DB 76; Length 868;
 Best Local Similarity 67.38; Pred. No. 2.4e-50;
 Matches 406; Conservative 0; Mismatches 161; Indels 36; Gaps 3;

QY 299 cgatccacatcagatcttcgcaagatagacgtgaacagtttccatcaacggttaactg 358
 DB 15 CGAAGCTAGATCATTTCTTAGAGAAATGAGTGAACAGATGTGCATCTACAGGCGGAT 74
 QY 359 cgaagagacacagaggttctcgtccaaacagatagatctcaagaccattagcgagaa 418
 DB 75 CGAAGAAAGAACTGAGATTTTCATCTCCGAACAGACATATTCAG---TGTGATGGGAA 131
 QY 419 gagaagtgaagaaagaaatctccggaaccggtctgctcggcgagatcaacgaga 478
 DB 132 AAGGAGCGAAAGGAGAGCACTAATGGA-----GATGA 164
 QY 479 gatactccgagatcagaggtactcagtggaacctcagatgaagaagaacgagggcga 538
 DB 165 GCTTGATATTGTAAGAGCTTGTCTCGTGCATCAGATGATGAGGAAAT-----GGTGA 218
 QY 539 aagtcgaggaagaagatcgtgtatcaaaagatcagcttcttctcgaagaagactt 598
 DB 219 CGCTTCGAGAAAGAAAGCTTAAAGCTTTCGAAGGATCAGTCTGCTATTCTTGAGAAAGCTT 278
 QY 599 caaagaacacacacatctcaatcccaacagaagctagcttggctaagaagctgaact 658
 DB 279 TAAAGAACACACACTCTGAAACCCAAACAAAGATGGCTTGGCTAAGCAGCTGGGATT 338
 QY 659 gaagcagaagacaagtggaagtgtgtccaaacagaagaagctgaacccaagttaagca 718
 DB 339 GCGACCCAGACAAAGTTAAGTTGTTGCCAAACAGAAAGGCCGAGGACCAAGCTGAAGCA 398
 QY 719 aacgagagtagatgcgaatacttgaacagtgctgtagaagaagctaaagaagacg 778
 DB 399 AACTGAGAGTTGACTGTGAGTTCTTAAAGACATGCTGTGACATATGACGAGGAAAAACAG 458
 QY 779 gagacttcagaagaagctatgagcttcgaactcgaactcgaactgtctccaaatctacgg 838
 DB 459 GCGGTTGAGAGAGAGAGTTCAAGAGCTGAGAGCAGCTGAACCTTCTCCTCAGTTCTACAT 518
 QY 839 tcagatgactccaccactacactcatcatgtctcctcgtgagaggttgaagctgtcc 898
 DB 519 GCAAAATACCCCACTTACGACCTCACCAGTGTGCCCATCGTGTGAGAGAGTAGGTGATC 578
 QY 899 atc 901
 DB 579 ATC 581

RESULT 12
 AM277753
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 ORGANISM
 soybean.
 Glycine max.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 607)
 REFERENCE
 AUTHORS
 Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corryell, V., Rhana
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuk
 R., Ritzer, E., Kohn, S., Shin, T., Jackson, I., Cadenas, M., McCann
 R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Public Soybean EST Project
 Public Soybean EST Project
 Contact: Shoemaker R/Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800


```

DB      242 TGCGGAGACATCTCATTACCAACA 212
RESULT  14
BE210250
LOCUS    473 bp          EST             21-NOV-2000
DEFINITION
soybean.
SOURCE    soybean.
ORGANISM  Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 473)
Shoemaker,R., Kelm,P., Vocklin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,D., Beck,C.,
Wyllie,T., Underwood,K., Stepien,M., Theisinger,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
TITLE     Public Soybean EST Project
JOURNAL   Contact: Shoemaker R/Public Soybean EST Project
COMMENT   Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel.: 314 286 1800
          Fax: 314 286 1810
          Email: est@wustl.edu
          This clone is available through: Genome Systems, Inc. 4633 World
          Parkway Circle St. Louis, Missouri 63134 For further information
          call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
          427-3324 or contact: clones@genomesystems.com or
          info@genomesystems.com web site: www.genomesystems.com
          Insert length: 675 Std Error: 0.00
          High quality sequence stop: 444.
FEATURES
Source
Location/Qualifiers
1..473
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl039-531"
/clone_lib="Gm-cl039"
/tissue_type="Whole seedling without cotyledons"
/lab_host="DH10B"
/note="Vector: pBluescriptII SK+; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from 2 week old seedlings with the cotyledons removed at
the time of harvest. The seedlings for the cultivar Oden
were grown in a growth chamber using germination paper.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into DH10B host cells (Gibco BRL). This
library was constructed by Dr. Randy Shoemaker."
BASE COUNT  131 a 131 c 121 g 89 t 1 others
ORIGIN
Query Match 18.1%; Score 221.2; DB 164; Length 473;
Best Local Similarity 73.3%; Pred. No. 7.5e-45;
Matches 283; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
```

OY	583	ttctcgaagagactctcaagaacaacacactctcaatcccaacaagaagctgcttg	641
Db	81	gtctcgggaagaagacttcaagbapcaaacactcttaaccgccacgaataattggac	140
OY	642	gctaaagaactgaacttgcgcgacaacagctggaagtctggtctccaaaacagaagact	701
Db	141	gggaacacactgggtcttgcagcgaacacaaagtgaagctgggttccagaaacaaaggca	200
OY	702	agaaccagttaaaacaaacggaagtagatctgcgaacttgaacggtgcgtagaagaag	761
Db	201	aggacgaacactgaagcaaacgaggttgactgcgaattttgaaaaggtctcgagaaat	260
OY	762	ctaagcgaagaaacccggagacttcaagaagaagcctatgagcttctgaactctcaagctg	821
Db	261	ctgacggagaaacacgaggttgccaggaagagtcacaaagcgtttgagagcgaataatt	320
OY	822	tctccacaactttagcttagatagatgacacacacactacactcaatctgctctcgtgc	881
Db	321	tccccgcactttctacttgacacattgaccccccacacaccccttcacactatgtccctctggt	380
OY	882	gagcgtgtagctggtctacatcatcatc	907
Db	381	gagcgtgtgcggtttcccccttcctc	406

RESULT	15
LOCUS	AM217402
DEFINITION	AM217402 550 bp mRNA EST
ACCESSION	EST296087 tomato flower buds 0-3 mm, Cornell University
VERSION	Lycopersicon esculentum cDNA clone cTGA1021 similar to Pimpinella
KEYWORDS	brachycarpa homobox-leucine zipper protein, mRNA sequence.
SOURCE	AM217402.1 GI:6528244
ORGANISM	EST.
REFERENCE	tomato.
AUTHORS	Lycopersicon esculentum
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
COMMENT	Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
	Lycopersicon.
	1 (bases 1 to 550)
	van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
	,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romning
	,C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
	Generation of ESTs from tomato flower tissue
	Unpublished (1999)
	Contact: David Frisch
	Clemson University Genomics Institute
	Clemson University
	100 Jordan Hall, Clemson, SC 29634, USA
	Tel: 864 656 4366
	Fax: 864 656 4293
	Email: dfrisch@clemson.edu
	5 prime sequence.
FEATURES	Location/Qualifiers
SOURCE	1..550
	/organism="Lycopersicon esculentum"
	/cultivar="TA496"
	/db_xref="taxon:4081"
	/clone="cTGA1021"
	/clone_lib="tomato flower buds 0-3 mm, Cornell University"
	/lssue_type="flower"
	/dev_stage="0-3mm buds"
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
	XhoI; supplier: Tanksley; Flower buds and flowers were
	taken from greenhouse plants (4-8 wks old, TA496). They
	were immediately frozen in liquid nitrogen and then
	size-separated while remaining frozen."
BASE COUNT	186 a 117 c 132 g 115 t
ORIGIN	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 10:45:17 ; Search time 7917.7 Seconds
(without alignments)
1875.599 Million cell updates/sec

Title: US-09-394-519-19

Perfect score: 1571

Sequence: 1 aggaacagtggaaggttcg.....gaagtttagttatgttgy 1571

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483.8	30.8	496	B29089	B29089 T26B14TR TA
2	417	26.5	417	B77441	B77441 T26B14TR TA
3	344.4	23.1	404	B27142	B27142 T29H6TR TAM
4	364	23.2	438	N65486	N65486 20526 Lambd
5	358	22.8	358	CNS00Y4	AL092818 Arabidops
6	354.6	22.6	758	B20632	B20632 T26E16-Sp6
7	313	19.9	436	H76458	H76458 18163 Lambd
8	301	19.2	304	F14268	F14268 ATTS5190 G1
9	301	19.2	487	AV53904	AV53904 AVS3904
10	168	10.7	231	CNS00M2C	AL093042 Arabidops
11	168	10.7	168	CNS00M2C	AL093042 Arabidops
12	165.4	10.5	213	Z34128	Z34128 ATTS3162 GT
13	145.2	9.2	763	B20629	B20629 T26B14-Sp6
14	138	8.8	410	BE524424	BE524424 M4969STM
15	119.6	7.6	492	BF79235	BF79235 GA_Eb003
16	85	5.4	486	BC596420	BC596420 EST495098
17	85	5.4	492	AM029972	AM029972 EST273227
18	85	5.4	494	AM033481	AM033481 EST277052
19	85	5.4	511	AM217152	AM217152 EST295866
20	85	5.4	513	AM623005	AM623005 EST320950
21	85	5.4	518	AM034705	AM034705 EST278507
22	85	5.4	605	AM625619	AM625619 EST319526
23	85	5.4	717	BC591978	BC591978 EST499820
24	81.8	5.2	462	AM979604	AM979604 EST341203
25	80.8	5.1	760	BC592817	BC592817 EST491595
26	79.2	5.0	510	BF050737	BF050737 EST435895
27	79.2	5.0	515	AM1896872	AM1896872 EST266315
28	79.2	5.0	545	AM18459	AM18459 EST320445
29	79.2	5.0	564	AT778378	AT778378 EST292527
30	79.2	5.0	564	AT778378	AT778378 EST292527
31	79.2	5.0	598	AM930485	AM930485 EST340858
32	79.2	5.0	601	AT895984	AT895984 EST265427
33	79.2	5.0	742	BG123412	BG123412 EST469058
34	77.6	4.9	561	AM035085	AM035085 EST279913
35	77.2	4.9	433	AM256333	AM256333 EST304540
36	77.2	4.9	478	BE247855	BE247855 NF022803D
37	77.2	4.9	558	BF520143	BF520143 EST457612
38	77.2	4.9	559	AM775766	AM775766 EST34831
39	77.2	4.9	612	BC583349	BC583349 EST485100
40	77.2	4.9	666	BC350434	BC350434 091E081Ma
41	77.2	4.9	681	BC598073	BC598073 EST496751
42	77.2	4.9	688	BF632128	BF632128 NF025D10D
43	77.2	4.9	541	AM032669	AM032669 EST276228
44	76.8	4.9	421	BE458391	BE458391 EST413683
45	76.8	4.9	503	AM096359	AM096359 EST289539

ALIGNMENTS

RESULT 1
B29089 496 bp DNA GSS 13-OCT-1997
LOCUS T26B14TR TAMU Arabidopsis thaliana genomic clone T26B14, DNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
TITLE
Use of a BAC End Sequence Database To Identify Minimal Overlaps for

JOURNAL COMMENT

Arabidopsis Genomic Sequencing
Unpublished (1997)
Other GSSs: T26B14TRC
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@igrr.org
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 496.
Location/Qualifiers
1. 496
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T26B14"
/clone_lib="TAMU"
/sex="hermaphrodite"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"

FEATURES

source

BASE COUNT

113 a 118 c 100 g 165 t

Query Match

Best Local Similarity 30.8%; Score 483.8; DB 256; Length 496;
Matches 485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	903	ttgatgtctatgttgcgtgaaatgtgtccactactaaacgtgatgttcttcatcgtaga	962
DB	1	TTGATGCTATGTGTCGGAATGTCCTACTACTAAACGTGATGTTCTTCATCTGAGA	60
QY	963	ctaaccaatgtcttcgttcttccactgttcttctgttggcaagaagacctctgcat	1022
DB	61	CTAACCAATGCTTCGTTCTTCCACTGTGTCCTGTTGACCAAGATGACACTTCTGCAT	120
QY	1023	cagcttcaacttgcgttcaacaacctgatgtacgtctgacccgttgcctcaactgtcca	1082
DB	121	CAGCTCACTTGTGTGTCACCAACCTGATGACGTCGACCCGTTGCTCCAACTGCTCAA	180
QY	1083	ctccaatgtcttcgttgcgttgcgaagaagaacgtgttgcattcgaattcactaattc	1142
DB	181	CTCCAAATGTCCTGCGTGTGGAACAAGAAAGCTTGTGATTTGCACTTACTAATC	240
QY	1143	tacagatccctgatttgggttcttgcgaagagagcgaagacctgacttgcattgtt	1202
DB	241	TACAGATCCCTGATTTGGTTTGGCGAGGAGCAACAGACCTTACCTTCATTTGTT	300
QY	1203	tcctcgagatgcacgttgcattgcgtctgctgctgattgcattcgaattcgaagattcga	1262
DB	301	TCCTCGGAGATGATCAGTTTATGATTTTCGCTGCTTGATGACATTCAGAGATTCGAG	360
QY	1263	ataacggttccagtcggttaccagattgcattgcgagatgttgaagattcgaactag	1322
DB	361	ATAACGGTCCAGTGCCTTACAGATTTGCACTTGGCGAGTGTGAAGATCTTCAGCTAG	420
QY	1323	ctgactctagtttgggttcttgcataactgtctctataaatctcttggccattaa	1382
DB	421	CTGACTCTAGTTTGGTTTCTTGCATCAACTGCTTCATCAACATCTTCCCATTTAA	480
QY	1383	aaatttt 1389	
DB	481	AAATTTT 487	

RESULT 2
B77441 417 bp DNA GSS 16-JAN-1998
LOCUS B77441
DEFINITION T26E16TR TAMU Arabidopsis thaliana genomic clone T26E16, DNA
sequence.

ACCESSION B77441
 VERSION B77441.1 GI:2774080
 KEYWORDS GSS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.
 REFERENCE 1 (bases 1 to 417)
 AUTHORS Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K., Berry,K., Granger,D., Sun,E., Whible,C., Adams,M.D. and Venter,J.C.
 TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3
 JOURNAL Unpublished (1997)
 COMMENT Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@tigr.org
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 417.
 Location/Qualifiers
 1. 417
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="T26E16"
 /clone_1ib="TAMU"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; site_1: HindIII; site_2: HindIII
 ; Produced by Rod Wing"
 BASE COUNT 96 a 96 c 90 g 135 t
 ORIGIN
 Query Match 26.5%; Score 417; DB 256; Length 417;
 Best Local Similarity 100.0%; Pred. No. 4.5e-95;
 Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 900 agtttgatgctatgtgtgtggaatgtgtccactactaaacgtgatttcttcacatg 959
 Db 1 AGTTGATGCTATGTTGCTGGAATGTGCGCACTACTAAAGCTGATGTTCTTCAATCG 60
 QY 960 agactagccaagtctgttcttccactgtgttccgtgttggaagaatgacattctg 1019
 Db 61 AGACTAGCCAAATGCTCTGCTTTCACCTGTTGCTCTGTTAGCAAGATGACACTTCTG 120
 QY 1020 catgaactcactgtgtgtcaacaaccctgatacgtctcgaacgttgcacaactgctc 1079
 Db 121 CATGACTCTCAGCTTGTGTCAACACCCGTGATGACGTCTCGACCGTTGCTCCAACTGCTC 180
 QY 1080 caactcacaatgttctctgtgtgtgaagaagaacgattgttctcgaatttacta 1139
 Db 181 CAACCTCAATATGTTCTGTGCGTGGGAACAAGAAAGTTGTCATTTGACTTTACTA 240
 QY 1140 atctacagatccctgatttctgttcttgcagagaagaacaagaactagactcgatt 1199
 Db 241 ACTTACAGATCCCTGATTTTGGTTCTTGTGCAAGAGACACTAGACTTGATTT 300
 QY 1200 gtttctcgcgagtgatcagttgattgattggtctgtctgtgtgacattcaagattcg 1259
 Db 301 GTTTCCTCGGGATGATCAGTTTGAATTCGGCTTGTGATGACATTCAGATTCG 360
 QY 1260 aagataacggtccaaagtgcgttacagatttcgaacttgcgagtggtgaagacttc 1316
 Db 361 AAGATAACGCTCCAAAGTGCCTTACAGATTTTCGACTTTGGGATGTTGAAGATCTTC 417
 RESULT 3

B27142 404 bp DNA GSS 13-OCT-1997
 LOCUS T29H6TR TAMU Arabidopsis thaliana genomic clone T29H6, DNA
 DEFINITION sequence.
 B27142
 ACCESSION B27142 GI:2513108
 VERSION B27142
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.
 REFERENCE 1 (bases 1 to 404)
 AUTHORS Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and Venter,J.C.
 TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing
 JOURNAL Unpublished (1997)
 COMMENT Other GSSs: T29H6TFB
 Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@tigr.org
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 404.
 Location/Qualifiers
 1. 404
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="T29H6"
 /clone_1ib="TAMU"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; site_1: HindIII; site_2: HindIII
 ; Produced by Rod Wing"
 BASE COUNT 91 a 98 c 83 g 132 t
 ORIGIN
 Query Match 25.1%; Score 394.4; DB 256; Length 404;
 Best Local Similarity 98.5%; Pred. No. 2.4e-89;
 Matches 398; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 895 gcttgagtttgatgcatgtgtgtgtgaatgtgtccactactaaacgtgatttcttc 954
 Db 1 GCTTGAGTTTGATGATGTTGCTGGAATGTGCTCACTACTAAACGTGATGTTCTTC 60
 QY 955 atctggaactagccaatgctcgttcttccactgtgttctcgtgttgagaagaatgac 1014
 Db 61 ATCTGAGACTAGCCAAATGCTCTGCTTTCACCTGTTGCTTCCGTGAGACATGACAC 120
 QY 1015 ttctgatacagctctcactgtgtgtcaacaaccctgtatgacgtctcgaacgttgc 1074
 Db 121 TTCTGCATCAGCTCTCAGTTGTGTCAACAACCTGATGAGTTCGACCGTTGCTCCAAC 180
 QY 1075 tgcctcaactccaatgttctcgtgtgtgtgaagaagaagaagttgttgcattcgact 1134
 Db 181 TGCTCCAACTCCAAATGTTCTGCTGTGGAACAACGAAACGTTGTTGATTTGCACTT 240
 QY 1135 tactaatcaagaatccctgatttctgttcttgcagagaaggagacaagaactagact 1194
 Db 241 TACTAATCTACAGATCCCTGATTTTGGTTCTTGTGCAAGAGACTACATACATTT 300
 QY 1195 cgattgttctcgcgagtgatcagtttgaatttcgaacttgcgattgcattgaacga 1254
 Db 301 CGATTGTTCTCGCGGATGATCAGTTTGAATTCGGCTTGTGATGACATTCGAATG 360
 QY 1255 atcgaagaataacggtccaaagtgcgttacagatttcgaacttgcattg 1298
 |||||

Db 361 ATTGCAAGATACGGTCCAGTGGCTTACGAGATTCCGACTTTG 404

RESULT 4
LOCUS N65486 438 bp mRNA EST 05-JAN-1998
DEFINITION 20526 Lambda-PRL2 Arabidopsis thaliana cDNA clone Z30D2177, mRNA sequence.
ACCESSION N65486
VERSION N65486
KEYWORDS EST.
SOURCE GI:1217112
ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 438)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E., and Somerville,C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9188
Email: 22313c@edlm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 438
/organism="Arabidopsis thaliana"
/strain="Var Columbia"
/db_xref="taxon:3702"
/clone="Z30D2177"
/note="Vector: lambda-Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."
BASE COUNT 99 a 90 c 98 g 134 t 17 others
ORIGIN

Query Match 23.2%; Score 364; DB 159; Length 438;
Best Local Similarity 90.1%; Pred. No. 1.2e-81;
Matches 391; Conservative 0; Mismatches 41; Indels 2; Gaps 1;

QY 790 ggggaaatggcgctgagatagatcctataagaanaactgagctgtgtgggtac 849
|||||
Db 1 GGGGAAATGGCGCTGAGATTAGATCCCTATTAAAGAACTGAGACTGTGGGTAC 60
|||||

QY 850 tttagaactcttgaagaagctgctaagctatgatgcttaagaagcttgatgac 909
|||||
Db 61 TTTTAAATCTTTGAAGAGGTGCTAAAGCTTATNATGCTAAGAGCTTGATTNATGC 120
|||||

QY 910 tatgttgcctggaatgctgctactactaaacgctatgttcttcaatctgagctagcca 969
|||||
Db 121 TATGTGCTCGTGAATATGCTCACACNACGATNATGTTCTTCATGCTGAGACTAGCA 180
|||||

QY 970 atgctctgctctcactcgtgtgtctgtgtgagcaagatgacactctgcatcagctct 1029
|||||
Db 181 ATGCTCTGCTCTTACACTGTTGTTGCCGTGAGCAAGATGACACTTCTGATCAGCTCT 240
|||||

QY 1030 cacttgtgtcaacaaccctgtatgacgtctcagccgtgtctccaaactgtccaaactcaaa 1089
|||||
Db 241 CACTTGTCTCAACAACCCGTATGACGTCTGACCCGTCTCCCAACNCTCCAACTNCAAA 300
|||||

QY 1090 tgttcctgctgtgtggaacaagaagctgtgtgatttgac--ttactaatctacag 1147
|||||
Db 301 TGTTCCTGCTGCTGGAACAAGAAACGTTGCTGCTTGGCTTACTAATNCTACAG 360
|||||

QY 1148 atccctgatttggcttctcctgagagagcaagaacactagacttcgatttctctc 1207
|||||
Db 361 TTCCTGTTTGGGTTTGTGGCAGAGGGGCCAACAGNCTTGGCCTTCGTTGTTCTCT 420
|||||

QY 1208 gcgagatgacagt 1221
|||
Db 421 NGGGGTNANCAAGTT 434
|||||

RESULT 5
LOCUS CNS00VW4/c 358 bp DNA GSS 28-JUN-1999
DEFINITION CNS00VW4 Arabidopsis thaliana genome survey sequence T7 end of BAC T10J21 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL092818
VERSION AL092818.1 GI:5293972
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 358)
Salanoubat,M., Choisy,N., Artiguenave,F., Brothier,P., Wincker,P., Samson,D., Saurin,M., Weissenbach,J. and Quetier,F.
Unpublished
2 (bases 1 to 358)
REFERENCE
Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191, 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Location/Qualifiers
1. 358
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T10J21"
/note="end : T7"
BASE COUNT 106 a 92 c 58 g 102 t
ORIGIN

Query Match 22.8%; Score 358; DB 219; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.9e-80;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 gaggaaagctccgtatctatgtgaatgacccctactgactgactgacattcccttaagtatga 594
|||||
Db 358 GAGGAAAGCTCCGTATTATTTGATGATGATCCCTTATGCTACTGATATCCCTCTAGGATGA 299
|||||

QY 595 ggaagagcttaagttctcctaagcaagaataaagctatgctgtgtgatttaact 654
|||||
Db 298 GGAAGAGCTTAAGGTTCTTAAGCCAAAGAAATGAAGATGATGCTGTGATTAACCTT 239
|||||

QY 655 tccctctatggaagcttctgacagcctctgagagcttctcctaaggaactactaaac 714
|||||
Db 238 TCCCTATATGGAATTTCTGAACAGCCTTCTGAGAGTTCTTCTCAGACATGACTAAAC 179
|||||

QY 715 tgatgcaagatagctgtgtcagctctcctgctgtgttccctaggaagaagcctgtgtgt 774
|||||
Db 178 TGATGCAAGATATACCTGTGCTAGCTTCTCGTGTTCTCTAGGAAGAAGCCTGTGTGTGT 119
|||||

OY 775 taaggcaagaataggaggagctgctgagattagagatctcattagaagaactag 834
 |||
 Db 118 TAGGCAAGAAATGGGGAAATGGCTGTGATTAAGATCTATTAAAGAACTAG 59
 |||
 OY 835 gacttggttggttactttgatatacttgaagaagctgtcaagcttaagatgaatga 892
 |||
 Db 58 GACTTGTTGGGTACTTTGATATCTCTTGAAGAAGCTGCTAAAGCTTATGATCTAG 1
 |||
 RESULT 6
 B20632 758 bp DNA GSS 16-SEP-1997
 LOCUS T26E16-Sp6 TAMU Arabidopsis thaliana genomic clone T26E16, DNA
 DEFINITION sequence.
 ACCESSION B20632
 VERSION B20632.1 GI:2395686
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 758)
 REFERENCE Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
 Ecker, J.
 BAC End Sequences at ATGC
 Unpublished (1997)
 CONTACT: Ecker J.
 Arabidopsis Thaliana Genome Center
 University of Pennsylvania
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA
 19104
 Tel: 215-898-9384
 Fax: 215-898-8780
 Email: jecker@atgenom.bio.upenn.edu
 Seq primer: Sp6
 Class: BAC ends
 High quality sequence start: 192
 High quality sequence stop: 293.
 FEATURES
 source
 1..758
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_1ib="TAMU"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
 ; Produced by Rod Wing"
 BASE COUNT 140 a 237 c 120 g 239 t 22 others
 ORIGIN
 Query Match 22.6%; Score 354.6; DB 256; Length 758;
 Best Local Similarity 86.7%; Pred. No. 3.4e-79;
 Matches 392; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
 OY 894 aggttaagtttgtatgtctgtgctggaatgtgtccactactaaccgtgatgttctt 953
 |||
 Db 75 AGCTTAGTGTGATGCTGATNGCTGGAATGTGTCACCTACTAAACGATGTTCTT 134
 |||
 OY 954 catctagagatagcaaatgctcgtctcttccactgttctgttgaagaagatgaca 1013
 |||
 Db 135 CATCTAGATATACCAATGCTCTGTTCTTACCTGTGNGTGTAGACCAAGATGACA 194
 |||
 OY 1014 ctctgcataagctctcaactgtgtcaacaaccctgatgaagctgcagcgtgtccaa 1073
 |||
 Db 195 CTCTCGATCAGCTCTACATTGTGTCAACAACCTGATGACGCTGACCGTTGCTCCA 254
 |||
 OY 1074 ctgctcaactcaaatgttctcgtcgtgtggaacaagaagcgtgttcagattcgcatt 1133
 |||

Db 255 CTGCTCCAACTCCAAATGTTCTGTGTGGAACAAGAAAGCTGTCCGATTTCGACT 314
 OY 1134 ttactaatctaacagatccctgatatttgcttctgtgcaagaagagcaaacagactagct 1193
 |||
 Db 315 TAAGTAATCTACAGATCCCTGATNTGGTCTGTGGGCAAGAGCAACAACTAAACT 374
 |||
 OY 1194 tccattgttccctcgcggatgatacagttgatattcgcgtcgttgatgacatcaag 1253
 |||
 Db 375 CCGATTGTNCCCTCGGGATGATGATCACTTGATGATTCGGCTTGAATCTCCNG 434
 |||
 OY 1254 gattcgaagataacggtccagatgcttacccagattcgaacttgcggatgtgaagato 1313
 |||
 Db 435 GATCCCAAAATNCCGG-CCAGTGCCTCACATATTCGACATTTGGCGGATGTCNATATC 493
 |||
 OY 1314 ttcagctagctgactctagtttcggttccct 1345
 |||
 Db 494 NNCCCTTTTCTTTTCTTTTCTTTTCTTTTCTT 525
 |||
 RESULT 7
 H76458 436 bp mRNA EST 05-JAN-1998
 LOCUS H76458
 DEFINITION 18163 Lambda-PRL2 Arabidopsis thaliana cDNA clone 195M2377, mRNA
 sequence.
 ACCESSION H76458
 VERSION H76458.1 GI:1053709
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 436)
 REFERENCE Newman, T., debruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh
 , L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel
 , E. and Somerville, C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcm@lhm.cl.msu.edu
 Seq primer: 5_prime.
 FEATURES
 source
 1..436
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone_1ib="Lambda-PRL2"
 /note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda zip-lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dT primed cDNA."
 BASE COUNT 109 a 80 c 107 g 124 t 16 others
 ORIGIN
 Query Match 19.9%; Score 313; DB 158; Length 436;
 Best Local Similarity 95.1%; Pred. No. 1e-66;
 Matches 331; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

OY 486 gagctaagaagaacatctatcgtccctcctcaagaaccacacattgaggaagtc 545
|||||
Db 1 GAGCTAAGAAGAACACTATCGTCCCTCCCTCAAGAAACCAACCTTGAGGAAGTCC 60
OY 546 gtattatgtgaatgatcctatgtactgtatgtatcctctagtgatgaagaagctta 605
|||||
Db 61 GTATTATGTGAATGATCTATGCTACTGATGATGATTCCTTAGATATAGGAAGAGCTTAA 120
OY 606 aggttcctcaagccaagaanaatgaacgtatcgttcgtgagattaacttcctctatg 665
|||||
Db 121 AGGTTCCCTCAACCAAGAAATGAACCTATCGTTCGAGATTATACCTTCTCTTATAGG 180
OY 666 aagttctgaacagcctctgagagttcctcgaagacagtaactaaactgagcaaga 725
|||||
Db 181 AAGTTCTGAACAGCCTTCTGAGAGTCTCTCAAGACAGACTAAACTGATGCAAGA 240
OY 726 tagctgtcagcctctcctcgtctgtcctctagaagaacccgtgtgttagg--caaag 783
|||||
Db 241 TAGCTGTCTGAGCTTCTCTCTCTCTCTAGGAACACCTGTTGCTTTAGGGCAAGG 300
OY 784 gaatgaggggaatggcgtgctgagatgagatcctatgaagaac 831
|||
Db 301 NAATGGGGGGAATGGCTTCTGAGCTTAGNGCTWCATTAAAGGAAC 348

RESULT 8
F14268/c 304 bp mRNA EST 27-JUN-1995
LOCUS F14268
DEFINITION ATTS5190 Gif-Seeda Arabidopsis thaliana cDNA clone YAY910 5', mRNA
sequence.
F14268
F14268.1 GI:871998
EST.
VERSION F14268.1
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 304)

REFERENCE
AUTHORS CNRS.
TITLE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Berthomieu P., Guerrier D., Giraudat J.
Genetique Moleculaire d'Arabidopsis
ISV - UPR40, CNRS
Avenue de la Terrasse, 1198 Gif-sur-Yvette Cedex, France
Email: Giraudat@cnrs-gif.fr.

FEATURES
source
1..304
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="YAY910"
/clone_lib="Gif-Seeda"
/note="Vector: Lambda ZAPII non-oriented; Physiological
condition: greenhouse plants. tissue_type: Green siliques"

BASE COUNT 108 a 65 c 47 g 82 t 2 others
ORIGIN

Query Match 19.2%; Score 301; DB 157; Length 304;
Best Local Similarity 99.3%; Pred. No. 1e-65;
Matches 301; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1252 aggttcgaagaatacagtcgaagtgcgtaccagattgcaacttgaggagtgtgaaga 1311
|||||
Db 304 AGGATTCGAATAACGGTCCAAAGCGCTTACCAAGATTTCACATTTGGGAGATGTTGAAGA 245
OY 1312 tcttcagctagcactagcttcttcgttccttgatcaactgctcctcatcaacatctc 1371
|||||
Db 244 TCTTCAGCTACTGACTGCTTCTGCTTCTGATCAACTTGTCTCTATCAACATCTC 185

OY 1372 ttgccattaaaagtttcgaagcttcataagatctgtcttagtaagtgaagagaag 1431
|||||
Db 184 TTGCCATTAAAAGTTTTCGAGCTTCATAGACATCTTCTTAGAATGTAAGTGAGAG 125
OY 1432 agtttctgttttcctttagtgccttagtaattaaagacatacaaaagtgtgttcgcg 1491
|||||
Db 124 AGTGTGTTGTTTTCGTTTATGCTTTAGTATTAATTMAACATACAAAAGTGtGtGTCG 65
OY 1492 gattctagaatcattagaacataaagccgggttttgcattaggaatcagatttcaat 1551
|||||
Db 64 GATTGCTAGTACATCTTAAGACATTAAGCCGGGTTTGCTTGAATTCAGATTTTAT 5
OY 1552 gaa 1554
|||
Db 4 GAA 2

RESULT 9
AV533904/c 487 bp mRNA EST 06-SEP-2000
LOCUS AV533904
DEFINITION AV533904 Arabidopsis thaliana flower buds Columbia Arabidopsis
thaliana cDNA clone FB070B02F 3', mRNA sequence.
AV533904
AV533904.1 GI:8694187
EST.
VERSION AV533904.1
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 487)

REFERENCE
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
source
1..487
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB070B02F"
/clone_lib="Arabidopsis thaliana flower buds Columbia"
/tissue_type="flower buds"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 196 a 94 c 74 g 123 t
ORIGIN

Query Match 19.2%; Score 301; DB 31; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.2e-65;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1271 ccaagtgctaccagattcgaacttgagagtgatgaagatcttcagctagctgactct 1330
|||||
Db 487 CCAAGTGGCTTACCAAGATTTCGACTTTCGGATGTTGAAGATCTGACGTGACTCT 428
OY 1331 agtttggttccttgatcaactgctcctcctaacaactcttcgccattaaagtttt 1390
|||||
Db 427 AGTTTGCTTTCCTTGACCAACTGCTCTCATCAACATCTTCCCACTTAAGAGTTT 368
OY 1391 gcaactcatagatccttgcttagtaagttaagtgaagaagtggtttgttttcggt 1450
|||||
Db 367 GCAGCTTCATAGGATCTTCCTTACTTAAGTAAAGTGAAGAGTGTTTGTTCGTT 308
OY 1451 tatgcttagtaattaaagacatacaaaagtgtgttcgagattgtagtaagatcttaa 1510

```

|||||
Db 307 TATGCTTATTAATTAAACATACAAAGTGTGTTCGGATGTTAGTAACATCTTAA 248
QY 1511 gacataagccgggttttgtaaataggacgcggttttaataagtttagtttagttt 1570
Db 247 GACATTAAGCGCGGTTTGCATTAAGATCGAGTTTAATGAAGTTTAAAGTTTATGTTT 188
QY 1571 g 1571
Db 187 g 187

RESULT 10
LOCUS CDS00W2C 231 bp DNA GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T10P3 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL093042 GI:5294196
VERSION AL093042
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 231)
REFERENCE Salanoubat,M., Choisme,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
AUTHORS Unpublished
JOURNAL 2 (bases 1 to 231)
REFERENCE Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
1. .231
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T10P3"
/note="end : T7"
BASE COUNT 66 a 63 c 37 g 65 t
ORIGIN

Query Match 14.7%; Score 231; DB 219; Length 231;
Best Local Similarity 100.0%; Pred. No. 5,3e-48;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 662 atggaagttctgaacagccttcctgagagttcttctcagacagctactaaactgagtc 721
Db 231 ATGGAAGTTTCTGAACGCCCTTCGAGAGTTCTTCAGACACTACTAAACTGAGTGGC 172
QY 722 aagatacgtgtcagcgttctcctcgtctcctcaggaagaagcctgtgtgttaagcaa 781
Db 171 AAGATACCTGTGCAGCTTCTCCTGCTGTTCTCAGGAAGAAGCCTGTGTTAGGCAA 112
QY 782 aggaatagggggaataggctgtcgtgagatagagatcctatlaagaanaactagacttg 841
Db 111 AGGAATAAGGGGAATGGCTGCTGAGATTAAGATCTTAAGAAACTAGAGACTTGG 52
QY 842 ttgggtactttgatactcttgaagaagcgtgctaaagcttatgatgctaag 892
Db 51 TTGGGTACTTTGTATCTTGAAGAAGCTGCTAAAGCTTATGATGCTAAG 1

RESULT 11
LOCUS CDS00TNP 168 bp DNA GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T4N22 of

```

```

TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL089923 GI:5291063
VERSION AL089923
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 168)
REFERENCE Salanoubat,M., Choisme,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
AUTHORS Unpublished
JOURNAL 2 (bases 1 to 168)
REFERENCE Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
1. .168
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T4N22"
/note="end : T7"
BASE COUNT 41 a 35 c 38 g 54 t
ORIGIN

Query Match 10.7%; Score 168; DB 219; Length 168;
Best Local Similarity 100.0%; Pred. No. 4,6e-32;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1128 tcgacttactaactacagatccctgatttggttcttcggcagaggaacaagacc 1187
Db 1 TCGACTTACTAATCTACAGATCCCTGATTTGTTCTTCGCGAGAGCAACAAGACC 60
QY 1188 tagactcgtgttctcctcgtcgagatcagtttgatgatttcggcttcgttgatgaca 1247
Db 61 TAGACTTGATGTTCTCTCGCGATGATCAGTTTGATGATTTGGCTTGATGACA 120
QY 1248 tccaagattcgaagaatacagtcaccaagtgcgtaccagatttcagact 1295
Db 121 TTCAGGATTCGAAGATACGGTCCAGTGCATTACCAAGTTTCAGCT 168

RESULT 12
LOCUS 234128 213 bp mRNA EST 14-JUN-1994
DEFINITION AF053162 Grenoble-B Arabidopsis thaliana cDNA clone GBE349 5', mRNA sequence.
ACCESSION 234128
VERSION 234128
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 213)
REFERENCE CNRS.
AUTHORS The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
CONTACT: Maché R., Quigley F., Thomas F., Yu DY.
COMMENT CNRS URA 1178
Laboratoire de Biologie Moléculaire végétale
B.P. 53X,38041 Grenoble Cedex,France
Email: FOURGLER@cicg990.grenet.fr.
FEATURES
source
1. .213

```

```

/organism="Arabidopsis thaliana"
/strain="ecotype Columbia C24"
/db.xref="taxon:3702"
/clone="G86e349"
/clone_1lb="Grenoble-B"
/note="Vector: Lambda ZAP; tissue=Flower buds of
A.thaliana ecotype Columbia C24; clone_library=Grenoble-B;
Cloning vector: Lambda ZAP"
BASE COUNT      47 a      30 c      52 g      83 t      1 others
ORIGIN

Query Match      10.5%; Score 165.4; DB 191; Length 213;
Best Local Similarity 87.4%; Pred. No. 2.2e-31;
Matches 181; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Oy 320 catagttgcgtgagaggttttgagtttgatatacaaaaaagtttbaaccttttagagttat 379
Db 6 CATAGTTGCGTGAAGAGTTTGTGATTTATATACAAAAAAGTTTACCTTTTAGAGTGAT 65

Oy 380 ttgttgcttcctcgtttcttcgtgatttttagagagtcggttaacatggttcagatt 439
Db 66 TTTTGTCTTCTTCCTTTCTTCGTGTTTTTTCAGAGAGCGGTTTAACATGGTTTCGATT 125

Oy 440 agaaagacagctctttagtggtgttagtagcagatataagaagagagcctaagaacac 499
Db 126 AGAAAGGAAAGACAGCTCTTTAGTGTGCTGTTGCGAATTTGGAGAGAGAGCTTAGAGAAC 185

Oy 500 acctatgcctccctcccaagaacc 526
Db 186 CTCCTTCGCTCCCTTCCTCCTCAGAGACC 212

RESULT 13
B20629 763 bp DNA GSS 16-SEP-1997
LOCUS T26B14-Sp6 TMMU Arabidopsis thaliana genomic clone T26B14, DNA
DEFINITION sequence.
ACCESSION B20629
VERSION B20629.1 GI:2395683
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eucots II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 763)
Feng,Y., Dewar,K., Buehler,E., Kim,C., Li,Y., Shum,P., Sun,H. and
Ecker,J.
BAC End Sequences at ATGC
Unpublished (1997)
Other_GSSs: T26B14-F7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 110
High quality sequence stop: 131.
Location/Qualifiers
1..763
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db.xref="taxon:3702"
/clone="T26B14"
/clone_1lb="TMMU"
/clone_2lb="hermaphrodite"
/note="Vector: BelovAC11; Site_1: HindIII; Site_2: HindIII"
FEATURES
Source

```

BASE COUNT	143 a	209 c	77 g	317 t	17 others
ORIGIN					
Query Match	9.2%	Score 145.2;	DB 256;	Length 763;	
Best Local Similarity	79.9%;	Pred. No. 3.9e-26;			
Matches 187;	Conservative 0;	Mismatches 45;	Indels 2;	Gaps 2;	
Oy	903	ttgatgtcatgtgtgct-ggaatgtgtccactactaactgtatgtttcttcacgcag	961		
Db	81	TTGAAGCATATNGTNGTCGGAAATCGCCACTACTAAACGTGATGTTCTTCATCTGAC	140		
Oy	962	actagccaatgtctcgcgtgttcacacgtgttctcgtgttgagcaagaatgacactctgca	1021		
Db	141	AGTACCCACCGCTCTCGTNCNTACACTGTNGTNCCTGTNGAGCAAGATGACACTTCTGCA	200		
Oy	1022	tcagctctcaactgtgtgtcaacaacccgtgaagtcgcgtcgcacgtgtgtccaaactgtcca	1081		
Db	201	TCAGCTCCTCACTTGTGTGTCACACCCCTGATGACGTCTGCAAGNNG-NCCTTCGNTTCA	259		
Oy	1082	actccaatgttctcgtcgtgtgtgagcaagaagaacgtgtgttcagtttcacattt	1135		
Db	260	ATTCCTATTTTTCGTTGCGGTATATGACACACTTGTTATGTATTATTTT	313		
RESULT 14					
BE524424					
LOCUS	BE524424	410 bp	mRNA	EST	19-MAR-2001
DEFINITION	M49G95TM Arabidopsis developing seed Arabidopsis thaliana cDNA				
ACCESSION	BE524424				
VERSION	BE524424.1	GI:9782402			
KEYWORDS	EST.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosid II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 410)				
AUTHORS	White,J.A., Todd,J., Newnan,T., Focks,N., Girke,T., Martinez de Ibarra,D.O., Javorski,J.G., Ohlrogge,J. and Benning,C.				
TITLE	A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil				
JOURNAL	Plant Physiol. 124 (4), 1582-1594 (2000)				
MEDLINE	20567808				
COMMENT	Contact: Benning, C				
	Dept. of Biochemistry & Molecular Biology				
	Michigan State University				
	224 Biochemistry, Michigan State University, East Lansing, MI 48824				
	, USA				
	Tel: 517 355 1609				
	Fax: 517 353 9334				
	Email: benning@msu.edu				
	Michigan State University DNA Sequencing Facility Arabidopsis Biological Resource Center, The Ohio State University, 309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX: 6142920603 TEL: 6142929371.				
FEATURES	Location/Qualifiers				
source	1. 410				
	/organism="Arabidopsis thaliana"				
	/strain="Columbia"				
	/db_xref="taxon:3702"				
	/clone="M49G9"				
	/clone_id="Arabidopsis developing seed"				
	/tissue_type="seed"				
	/dev_stage="5-13 days after flowering"				
	/lab_host="E. coli"				
	/note="Organ: Developing seed; Vector: pBluescript SK-;				
	Site_1: EcorI, Site_2: XhoII"				
BASE COUNT	97 a	83 c	88 g	142 t	
ORIGIN					

Query Match 8.8%; Score 138; DB 136; Length 410;
 Best Local Similarity 66.8%; Pred. No. 2.2e-24;
 Matches 270; Conservative 0; Mismatches 105; Indels 29; Gaps 4;

QY 1 aagacagtgaaaggttcggttttgggttcgatacacaagaagaagg 60
 |||||
 DB 11 AGGACAGTAAAGTCGGGATTTTGGGATTTGAGAACATCGACACACGGCG 70
 |||||
 QY 61 ttt-----gattatcggcgtggttgatcagctgatttcttgatca 112
 |||||
 DB 71 TTTCGTACCGGGATTTATGCGGCTGTGTGAAATCGGCGCATTTGCTGTTC 130
 |||||
 QY 113 tatctcttcgcgttcatacctcctccatccatcgctgcttctgaatcgtc- 171
 |||||
 DB 131 AATCTCTTTCGAAATTCANCGCTCTTCTTCTCTGCTGTGTGTTTGTGTG 190
 |||||
 QY 172 -----ttctcaagctctcactctgctgtaataagaaggcttct 217
 |||||
 DB 191 GAGTCTGACGGGCTGTCTTAACGCTCTGCTCTGCTCTCTCTCTCTCT 247
 |||||
 QY 218 gagagactctctcttccttcacatgagcttaagaccacaaggactgttctagttgaagt 277
 |||||
 DB 248 TGAGGTTCTCTCCATTTCCATGGGCTTAAGACTTAAAGGACTGTCTGTGTTGAAGT 307
 |||||
 QY 278 ctctgggggtttcacataagaagcaaaagtctcttcttcataagctcgtagagatt 337
 |||||
 DB 308 ATTGGGGGTTTTCACGAAAGCAAAAAA---TACCTTTTTCATAGTTGCTGGAACAG 364
 |||||
 QY 338 ttgagtttgatacacaagaagtttgaccttttagagtgattt 381
 |||||
 DB 365 AGAAATCTTTTCTGTAAGAAACATCCGATTTAGCAGTTT 408
 |||||

RESULT 15

BF279235 492 bp mRNA EST 07-MAR-2001
 LOCUS GA_Eb0037N14f Gossypium arboreum 7-10 dpa fiber library Gossypium
 DEFINITION arboreum cDNA clone GA_Eb0037N14f, mRNA sequence.

ACCESSION BF279235
 VERSION BF279235.2 GI:13248937

KEYWORDS EST.
 SOURCE Gossypium arboreum.

ORGANISM Gossypium arboreum.

REFERENCE Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliopsida: eudicotyledons: core eudicots;
 Rosidae: eudicots II; Malvales; Malvaceae; Gossypium.

AUTHORS Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
 1 (bases 1 to 492)

TITLE An integrated analysis of the genetics, development, and evolution
 of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT On Nov 17, 2000 this sequence version replaced gi:11210305.

CONTACT: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu

Seq primer: TAATAGACTCACTATAGCG
 High quality sequence stop: 481.

FEATURES
 SOURCE 1..492
 Location/Qualifiers

1..492
 /organism="Gossypium arboreum"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_Eb0037N14f"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /tissue_type="Fibers isolated from bolls harvested 7-10
 dpa"

/lab_host="E. coli"
 /note="Vector: pBR-CMW; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 142 a 117 c 109 g 124 t
 ORIGIN

Query Match 7.6%; Score 119.6; DB 146; Length 492;
 Best Local Similarity 73.8%; Pred. No. 1.1e-19;
 Matches 152; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 741 ctccctgtcttctcaggaagagccctgtgtgtgttgagcaaggaatgagggaatgg 800
 |||||
 DB 69 CTCGTGTGGCTTAAMACCAAAAAACCTGTGGGCTTCTGTAAGAGAGTGCGCAATGGG 128
 |||||
 QY 801 ctgctgagattagagatcctatttaagaagactagagcttggttggttactttgatactc 860
 |||||
 DB 129 CTGCTAGATCAAGCATTCACCTCAAGAAAAACGAGATTGGTTGGAATTTACGATACCC 188
 |||||
 QY 861 ttgaagaagctgctaagcttaltgactaagaagcttgattgactatgtgtcgt 920
 |||||
 DB 189 TTGAAGAAGCAGCAAGGCTTATGAAGCTAAGAGCTGAGTTTGAAGCTGCTCATCTGTG 248
 |||||
 QY 921 gaaatgtgtccactactaaagtgtat 946
 |||||
 DB 249 CTGCCCGCCGCTCTGCTCATCTGAT 274
 |||||

Search completed: June 19, 2001, 17:00:23
 Job time: 22506 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:46 ; Search time 7917.7 Seconds
(without alignments)
1362.227 Million cell updates/sec

Title: US-09-394-519-123

Perfect score: 1141
Sequence: 1 aggaataaataacatgaty.....gtacacaaaaaaaaaaaaa 1141

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq-length: 0

Maximum DB seq-length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_estl48:*
118: gb_estl49:*
119: gb_estl50:*
120: gb_estl51:*
121: gb_estl52:*
122: gb_estl53:*
123: gb_estl54:*
124: gb_estl55:*
125: gb_estl56:*
126: gb_estl57:*
127: gb_estl58:*
128: gb_estl59:*
129: gb_estl60:*
130: gb_estl61:*
131: gb_estl62:*
132: gb_estl63:*
133: gb_estl64:*
134: gb_estl65:*
135: gb_estl66:*
136: gb_estl67:*
137: gb_estl68:*
138: gb_estl69:*
139: gb_estl70:*
140: gb_estl71:*
141: gb_estl72:*
142: gb_estl73:*
143: gb_estl74:*
144: gb_estl75:*
145: gb_estl76:*
146: gb_estl77:*
147: gb_estl78:*
148: gb_estl79:*
149: gb_estl80:*
150: gb_estl81:*
151: gb_estl82:*
152: gb_estl83:*
153: gb_estl84:*
154: gb_estl85:*
155: gb_estl86:*
156: gb_estl87:*
157: gb_estl88:*
158: gb_estl89:*
159: gb_estl90:*
160: gb_estl91:*
161: gb_estl92:*
162: gb_estl93:*
163: gb_estl94:*
164: gb_estl95:*
165: gb_estl96:*
166: gb_estl97:*
167: gb_estl98:*
168: gb_estl99:*
169: gb_estl00:*
170: gb_estl01:*
171: gb_estl02:*
172: gb_estl03:*
173: gb_estl04:*
174: gb_estl05:*
175: gb_estl06:*
176: gb_estl07:*
177: gb_estl08:*
178: gb_estl09:*
179: gb_estl10:*
180: gb_estl11:*
181: gb_estl12:*
182: gb_estl13:*
183: gb_estl14:*
184: gb_estl15:*
185: gb_estl16:*
186: gb_estl17:*
187: gb_estl18:*
188: gb_estl19:*
189: gb_estl20:*

190: gb_estl21:*
191: gb_estl22:*
192: gb_estl23:*
193: gb_estl24:*
194: gb_estl25:*
195: gb_estl26:*
196: gb_estl27:*
197: gb_estl28:*
198: gb_estl29:*
199: gb_estl30:*
200: gb_estl31:*
201: gb_estl32:*
202: gb_estl33:*
203: gb_estl34:*
204: gb_estl35:*
205: gb_estl36:*
206: gb_estl37:*
207: gb_estl38:*
208: gb_estl39:*
209: gb_estl40:*
210: gb_estl41:*
211: gb_estl42:*
212: gb_estl43:*
213: gb_estl44:*
214: gb_estl45:*
215: gb_estl46:*
216: gb_estl47:*
217: gb_estl48:*
218: gb_estl49:*
219: gb_estl50:*
220: gb_estl51:*
221: gb_estl52:*
222: gb_estl53:*
223: gb_estl54:*
224: gb_estl55:*
225: gb_estl56:*
226: gb_estl57:*
227: gb_estl58:*
228: gb_estl59:*
229: gb_estl60:*
230: gb_estl61:*
231: gb_estl62:*
232: gb_estl63:*
233: gb_estl64:*
234: gb_estl65:*
235: gb_estl66:*
236: gb_estl67:*
237: gb_estl68:*
238: gb_estl69:*
239: gb_estl70:*
240: gb_estl71:*
241: gb_estl72:*
242: gb_estl73:*
243: gb_estl74:*
244: gb_estl75:*
245: gb_estl76:*
246: gb_estl77:*
247: gb_estl78:*
248: gb_estl79:*
249: gb_estl80:*
250: gb_estl81:*
251: gb_estl82:*
252: gb_estl83:*
253: gb_estl84:*
254: gb_estl85:*
255: gb_estl86:*
256: gb_estl87:*
257: gb_estl88:*
258: gb_estl89:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

JOURNAL MEDLINE COMMENT		size-selected cDNA libraries DNA Res. 7, 175-180 (2000) 20363093		Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 153-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.	
FEATURES		Source		location/Qualifiers	
BASE COUNT		136 a 107 c 89 g 111 t			
ORIGIN					
Query Match		38.8%; Score 443; DB 31; Length 443;			
Best Local Similarity		100.0%; Pred. No. 6.9e-110;			
Matches 443; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
QY	681	tgcgaactcttgaaacggtgcgtatgagaagctaacggaaggaacggaagcttcagaa	740		
DB	443	TGCGAATACCTTGAAACGGTGGTAGAGAGCTAACGGAAGAACCGGAGACTTCAGAA	384		
QY	741	gagcctatgagcttcgaactccaagctgtctccacaattctacggtcagatgctcca	800		
DB	383	GAGGCTATGGAGCTTCCAGACTCTCAAGCTGTCTCCACAATTCTACGGTCAGATGACTCCA	324		
QY	801	ccaactcacactcatalgtgtccttcgtgcgagcggtgtgggtggccacatcaltcgaaac	860		
DB	323	CCAACTACACTCATTGATGTCGCTTCGCGAGCGGCGTGGGGCCCATCATCATCGAAC	264		
QY	861	cataccacaactacaagggccggtttctatacaatccgtyggtgtgctgtgctgaagt	920		
DB	263	CATCACACAAATCACAGAGCCCGTTTCATCAATCCGTGGTTCCTTGCTGCTCAGGTG	204		
QY	921	gtcatggtgcgaattttgaagccttggtccaagatcgatatttttatattagtggt	980		
DB	203	GCTCATGGGCTGAATTTTGAAGCCCTGCGTCCAGCAGCTGATTTTATTTTAACTGAGT	144		
QY	981	ggaagaaggtgtttgtgtatttcgtatcgtatatactatagctatctgtgtggggcatt	1040		
DB	143	GGAAAAAGGCTGTTTGGGATTTTTCGTTATGCTATATAGCTATCTGTGGGGTCAATT	84		
QY	1041	gtaatttgatgatctgagccttcctcatgaactagtcctatgatgatgacaacttaaaa	1100		
DB	83	GTAATTTTGGATGATGTTGGCCCTTCATGACACTAGTCATATGATGATGATCAACCTTA	24		
QY	1101	gatttaaatgacaaaattagct 1123			
DB	23	GATTTAAATTAGCAAAATTAAGT 1			
RESULT 2					
LOCUS		A1100243			
DEFINITION		545 bp mRNA		EST	
ACCESSION		A1100243		21-AUG-1998	
VERSION		A1100243.1		GI:3450204	
SOURCE		EST.			
ORGANISM		thale cress.			
		Arabidopsis thaliana			
		Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
		Rosidae; eucotids II; Brassicales; Brassicaceae; Arabidopsids.			
		1 (bases 1 to 545)			

AUTHORS	
Nemman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlirrogge, J., Ralkehl, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C.	
TITLE	
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones	
JOURNAL	
Plant Physiol. 106, 1241-1255 (1994)	
MEDLINE	
95148729	
COMMENT	
Contact: Thomas Nemman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel: 517-353-0854 Fax: 517-353-9168 Email: 22313ycn@lm.cl.msu.edu The sequence entry for this EST has been reverse complimented and is being submitted in the sense orientation. Seq primer: M13-Universal. Location/Qualifiers 1. 545 /organism="Arabidopsis thaliana" /strain="var columbia" /db_xref="taxon:3702" /clone="110C16XP" /clone_id="Lambda-PRL2" /note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not; lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark - rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."	
BASE COUNT	
129 a 116 c 131 g 163 t 6 others	
ORIGIN	
Query Match	36.6%; Score 418; DB 16; Length 545;
Best Local Similarity	96.4%; Pred. No. 4.7e-103;
Matches 449; Conservative 0; Mismatches 12; Indels 5; Gaps 2;	
01	taaacgaacgaaggttaattgctgaataacggttgcttaagaagcgaacgaag 720
02	
03	85 TTTAANNAACGGAGGTAGATT---CGATTGAACGGGTCTAGAGAAGCTTAACGGAAG 140
04	
05	721 agaaccggaagactcaagaagaagagctatgagagcttcgaactccaagctgtctccacaat 780
06	
07	141 AGACCCGGAGACTCCAG--AAGAGGCTATGAGACTTCGAACCTCTCAAGCTGCTCCACAAT 199
08	
09	01 tctaaagtcagatgactccaccaactaacatcatcatatgtctcttcgagagcgttgg 840
10	
11	200 TCTAGCGTCAGATGACTCCACCACTAACATCATATCATGTCTTCGCGCGAGCTGGG 259
12	
13	841 gtggccatcatcatcagcaaccacacacacacacacagcggcttcatacactcgtgg 900
14	
15	260 GTGGCCCATCATCATCGAAGCAACATCAACCAATCAAGAGCCGCTTCTATCATCCGTGG 319
16	
17	01 ttgcttctgctgagtcagtgagctcatgagctgaatttgaagccttgcgtccagatcgt 960
18	
19	320 TTGCTTTGCTGCTGTCAGGTGCTCATGTGGCTGAATTTGGAACCTTTGGCTCCACATCGT 379
20	
21	01 gatatttatttaattggtgggaagaagggtgtttggtatttttgatatgatatag 1020
22	
23	360 GATTTTATTTTATTTAATGCTGGGAAAAAGGCTTTTGGTATTTTTCGTTATCGTATATAG 439
24	
25	01 tctatctgtgtgggtcattgtcaattttgagatgattgccccttcataacacagcctat 1080
26	
27	440 TCTATCTGTGTGGGTCAATGTATTTTGGATGATTTGGCCTTCTCATGAACTAGTCCAT 499
28	
29	01 gtaagagcaaccttaaaaagatttaattgaattagcaaaattagttac 1126
30	

Db	500	GTATGATGCACCTTAAAAAGATTAAATTAACGNMAAATAACTTAC	545
RESULT	3		
LOCUS	AV556109/c		
DEFINITION	AV556109 Arabidopsis thaliana green siliques Columbia Arabidopsis thaliana cDNA clone S0034g02F 3', mRNA sequence.		
ACCESSION	AV556109		
VERSION	AV556109.1		
KEYWORDS	GI:8727524		
SOURCE	EST.		
ORGANISM	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized antisense-selected cDNA libraries		
AUTHORS	DNA Res. 7, 175-180 (2000)		
TITLE	20363093		
JOURNAL	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamiz@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/location/Qualifiers		
MEDLINE	Location/Qualifiers		
COMMENT	1..375		
FEATURES	/organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="S0034g02F" /clone_1fb="Arabidopsis thaliana green siliques Columbia" /tissue_type="green siliques" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	126 a 84 c 78 g 87 t		
ORIGIN			
Query Match	32.7%; Score 373.4; DB 31; Length 375;		
Best Local Similarity	99.7%; Pred. No. 6e-91; Mismatches 1; Indels 0; Gaps 0;		
Matches 374:	Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	752 gcttcgaactcctaagctgtctccacaattctacggctcagaatgactccaactaacact	811	
Db	375 GCTTGGAACCTCAAGCTGTCTCCACAATTCTACGGTCAGATGACCTCCACAATCTACT	316	
Qy	812 catcttgctcttcgttgtagcgctgtgtgtgtgcaccatcatatcgaaacctcacacaa	871	
Db	315 CATCATGTCCTCTGCTGGAGCGTGTGGTGGCCCATCATCATCAACACATCACACAAA	256	
Qy	872 tcacaggcccgtttctatcaatcacgcgtgtgtgtctgtctgtgtgtgtgtgtgtgtgt	931	
Db	255 TCACAGGCCCGTTTCTATCATATCCGIGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	196	
Qy	932 gaattttgaagccctgcgtgccacgatcgtgatttttatatttaagtgtgtgaaaaggtg	991	
Db	195 GAATTTTGAAGCCTTGCGTCCAGCAACGCGATTTTTTATTTTAGTGTGGGAAAAGGGTG	136	
Qy	992 ttgttgatatcttcgtctatcgtaataatgctatcatcgtgtgtgtgtgtgtgtgtgtgt	1051	
Db	135 TTTTGTATTTTTTCGTTATCGTTATATAGTCATATCTGTGTGTGGGTCATTTGTATTTTGA	76	
Qy	1052 tgattggcccttcacgaactagctcatatgatgatgatgacacacttaaaaaagatttaata	1111	
Db	75 TGATTGGCCTTCGATGAATAGTCTTATGTATGATGATGATGATGATGATGATGATGATGAT	16	
Qy	1112 gcaadaaatagttac 1126		
Db	15 GCAAAAAATTAGTTAC 1		

```

RESULT 4
LOCUS T42030 475 bp mRNA EST 07-JAN-1998
DEFINITION 5293 Lambda-PRL2 Arabidopsis thaliana cDNA clone 110C16T7, mRNA
sequence.
ACCESSION T42030
VERSION T42030.1 GI:932917
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
COMMENT On Nov 29, 1993 this sequence version replaced gi:634618.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lpm.cl.msu.edu
Seq primer: T7.

FEATURES
source
1..475
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="110C16T7"
/note="Vector: lambda Zip-Lox. Site_1: Sal. Site_2: Not.
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT 144 a 99 c 111 g 103 t 18 others
ORIGIN
Query Match 31.8%; Score 363.4; DB 188; Length 475;
Best Local Similarity 88.2%; Pred. No. 3.3e-88;
Matches 419; Conservative 0; Mismatches 49; Indels 7; Gaps 3;

QY 1 aggaataaaacatgatgatggaagaagaatcaggttgagcctaagttaggt 60
|
|
|
Db 1 AGGAATAAATAAACATGATGATGATGGCAAGAAAGATCTAGGTTGAGCCTAAGCTTAAGGT 60
|
|
|
QY 61 ttccacaatacacaatcctcttcagatgatcgaatcctaactctcatcattcaaca 120
|
|
|
Db 61 TTTCACAAATTCACAAATCCTCTTCAGATGATGATCTAATCTTCACTATATCAACA 120
|
|
|
QY 121 atctcagaagactccatcgaagaaacaaatcgaatcctcatcagatcgcagaatag 180
|
|
|
Db 121 ATCTCAGAGACTCCCATGGAACCAACATTCATCCTACATCAGATCTTCGCAAGATAG 180
|
|
|
QY 181 acgtgacaagcttccatcaacgggttaactgaggaagaacacagaggttcgcacaa 240
|
|
|
Db 181 ACGTGAACAGCTTTCATCAACGGTTTAACTGCGAGAGACACAGAGTTTCGTACCAA 240

```

```

QY 241 acagtagatctcaagcaccaattagcgggaagaagtgagagagaatctccgaa 300
|
|
|
Db 241 ACAGTAGATCTCAAGCACCAATTAGCGGGAAGAAGATGAGAGAAATCTCCGAA 300
|
|
|
QY 301 ccggcgcttgctccgggaagatcagacagatcattccgcgagggg--tactaacg 358
|
|
|
Db 301 CCGGCGTTGGCTCCGGGACGNTCAGCGGATCCTCGGATCGAGGGCTACTTACG 360
|
|
|
QY 359 tgaacctcagat---gaagaagaagcgggagcgaacgctcgaggaagaagtcaggt 414
|
|
|
Db 361 TGAACCTCAGATTTGAGAGAGAACGCGGCGGACGTCGNNNGGAAGAANTNNGT 420
|
|
|
QY 415 tatcaaatcagctcgtcttcttcgaagagacttcaagaacacacactct 469
|
|
|
Db 421 TTTTCAAGAATCGNCCTCTCTTTTGTG-GGGGNTTTCANNGNCAAAANNTT 474
|
|
|
RESULT 5
AV556091/c 372 bp mRNA EST 06-SEP-2000
LOCUS AV556091 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION thaliana cDNA clone S0034C01F 3', mRNA sequence.
ACCESSION AV556091
VERSION AV556091.1 GI:8727506
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
1 (bases 1 to 372)
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..372
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="S0034C01F"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 122 a 91 c 78 g 81 t
ORIGIN
Query Match 31.2%; Score 356; DB 31; Length 372;
Best Local Similarity 97.3%; Pred. No. 3.3e-86;
Matches 362; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 731 acttcagaagagctatgagcttcgaactcgaactcgaactcgaactcgaactcga 790
|
|
|
Db 372 ACTTCAGAAAAGAGCTATGAGAGTTTGACCTCCACACCTGCTCCACAGGTTTACGCTCA 313
|
|
|
QY 791 gatgactccacaactacatcattcattccttgctgtagagcgggtggtggccatc 850
|
|
|
Db 312 GATGACTCCACCAACTACATCATCATGATGATGCTTCCTGCGACCGTGTGGTGGCCATC 253
|
|
|
QY 851 atatcgaacatcacacacatacacaagcggcttcctatcaacccgtgggttgctgtgc 910
|
|
|
Db 252 ATATGAAACATCCCAACATACAGAGCCCGTTTCTATCAATCCGTGGTGTGCTGTGC 193

```

QY 911 tggtagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 970
 |||||
 Db 192 tggtagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 133
 |||||
 QY 971 tttagtgctgaggaagagctgtttgtatcttgcgtatctatatagtctatctgtg 1030
 |||||
 Db 132 tttagtgctgaggaagagctgtttgtatcttgcgtatctatatagtctatctgtg 73
 |||||
 QY 1031 tggtagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 1090
 |||||
 Db 72 tggtagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 13
 |||||
 QY 1091 accttaaaaaa 1102
 |||||
 Db 12 accttaaaaaa 1
 |||||

RESULT 6
 AI994446 526 bp mRNA EST 08-SEP-1999
 LOCUS 701498340 A. thaliana, Ohio State clone set Arabidopsis thaliana
 DEFINITION CDNA clone 701498340, mRNA sequence.
 ACCESSION AI994446
 VERSION AI994446.1 GI:5841351
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 526)
 REFERENCE
 AUTHORS Chen, J., Momiyama, M., Chan, E., Mooney, M., Carreon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobriza, A., Murry, L., Turner, C., Kikorian, S., Elder, L. and Hanson, D.
 TITLE Arabidopsis thaliana Gene Expression Microarray
 JOURNAL Unpublished (1999)
 COMMENT Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 463 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES
 source
 1. 526
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="701498340"
 /note="CDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."
 BASE COUNT 170 a 122 c 103 g 130 t 1 others
 ORIGIN

Query Match 29.7%; Score 338.4; DB 104; Length 526;
 Best Local Similarity 96.5%; Pred. No. 2.2e-81;
 Matches 356; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 aggaataataaacaatgatgatgggaagaagatctagattgagcctaagcttgggt 60
 |||||
 Db 148 aggaataataaacaatgatgatgggaagaagatctagattgagcctaagcttgggt 207
 |||||
 QY 61 ttccacaaatcaacaatcctcttcagatgaatcgaatcctaactctatatacaaca 120
 |||||
 Db 208 ttccacaaatcaacaatcctcttcagatgaatcgaatcctaactctatatacaaca 267
 |||||
 QY 121 atctccagagactccatggaacaaacatctgatctcatcagatcttcgcaagatg 180
 |||||
 Db 268 atctccagagactccatggaacaaacatctgatctcatcagatcttcgcaagatg 327
 |||||

QY 181 acgttaacagctttccatcaacggttaactgcgaggaacacagaggttcgtccaca 240
 |||||
 Db 328 acgttaacagctttccatcaacggttaactgcgaggaacacagaggttcgtccaca 387
 |||||
 QY 241 acaagtaacagcttcac 300
 |||||
 Db 388 acagtaacagcttcac 447
 |||||
 QY 301 ccggagtgctg-cctccgagacacacacacacacacacacacacacacacacac 359
 |||||
 Db 448 acggcgttgctg-cctccgagacacacacacacacacacacacacacacacacac 507
 |||||
 QY 360 ggaacctca 368
 |||||
 Db 508 gcnacctca 516
 |||||

RESULT 7
 H76307 390 bp mRNA EST 30-DEC-1997
 LOCUS 18012 Lambda-PRL2 Arabidopsis thaliana CDNA clone 200B97T, mRNA
 DEFINITION sequence.
 ACCESSION H76307
 VERSION H76307.1 GI:1053558
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 390)
 REFERENCE
 AUTHORS Newman, T., deBrujn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, J. and Somerville, C.
 TITLE Geiges galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis CDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tneibm.cl.msu.edu
 Seq primer: 17 dye primer.

FEATURES
 source
 1. 390
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="200B97"
 /note="Vector: lambda ZAP-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a CDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda ZAP-Lox. The CDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed CDNA."
 BASE COUNT 94 a 79 c 95 g 107 t 15 others
 ORIGIN

Query Match 27.4%; Score 312.4; DB 158; Length 390;
 Best Local Similarity 90.4%; Pred. No. 2.5e-74;
 Matches 347; Conservative 0; Mismatches 35; Indels 2; Gaps 2;

```

OY 654 accaagttaagaacacgaggtagatgtcgaacttgaaacggtgcgtagagaacta 713
    |||||||
DB 1 ACCAAGTTAAAGCAACGGAGGTAGATTGCAAACTTGAACGGTCTGAGAGAACTA 60
OY 714 acggaagaacacggaactcagaagaagctatagagcttcgaactccaagctctc 773
    |||||||
DB 61 ACGGAANANACGGGACACTTCAGAAAGGCGATGAGAGCTTGAACCTCAGACTGTCT 120
OY 774 ccaaaattcagtcagatgactcccaactaactcatcatcgttccctcgtcgaag 833
    |||||||
DB 121 CCACAATTNNCGGTGATGATGACTCCACCAACTACACTCATCATGNTTCCCTGCGGAG 180
OY 834 cgtgtgtgtgtcccatcatcatcatgaaccatcaaccacaatacagccgtttcataat 893
    |||||||
DB 181 CMTNTGGGTGGCCCATCATCATGCAACCAACCAACCAACCAACCAACCAACCAACCAAT 240
OY 894 ccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 952
    |||||||
DB 241 CCGTGGGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 300
OY 953 acgatcgtga-tttttatttttagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1011
    |||||||
DB 301 ACGTGTCAAAATTTTNTTNTTAAAGTGTGGGAAAGGCTGTTTGGCATTTTTCGGNATC 360
OY 1012 gtatatagctctatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1035
    |||||||
DB 361 GNAATATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 384

RESULT 8
LOCUS AV556640 276 bp mRNA EST 06-SEP-2000
DEFINITION AV556640 Arabidopsis thaliana green siliques Columbia Arabidopsis
ACCESSION AV556640
VERSION AV556640.1 GI:8728055
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 276)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL 20363093
MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/
FEATURES
source
1. 276
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="S0048f05F"
/clone_1ib="Arabidopsis thaliana green siliques Columbia"
/tissue-type="green siliques"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 97 a 68 c 50 g 61 t
ORIGIN
Query Match 23.6%; Score 269.6; DB 31; Length 276;
Best Local Similarity 98.6%; Pred. No. 1e-62;
Matches 272; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 836 tctgggtgagccatcatcatcgaaccatcacacaatcacagccgtttctatcaatcc 895

```

```

DB 276 TGTGGTGGCCCATCATCATGAAACATCCCAACATCACAGGCCGCTTCTATCAATCC 217
OY 896 gtgggtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 955
    |||||||
DB 216 GTGGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157
OY 956 atcgtatcttctatcttgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1015
    |||||||
DB 156 ATCGGATTTTATTTTATTTAGTGTGGGAAAGGCGTGTGTGTGTGTGTGTGTGTGTGT 97
OY 1016 tatagtctatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1075
    |||||||
DB 96 TATAGTATATCATGAGTGGGCTCATTTGATTTTGGATGATGAGGCGCTTCATGAACTAGT 37
DB 36 CCAATGATATGATGCAACCTTAAAGATTAAATTA 1

RESULT 9
LOCUS AI099721 389 bp mRNA EST 21-AUG-1998
DEFINITION 33874 Lambda-PRL2 Arabidopsis thaliana cDNA clone 120N14XP 3', mRNA
sequence.
ACCESSION AI099721
VERSION AI099721.1 GI:3449460
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 389)
Newman, T., deBrujn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh
, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomasnow, M., Retzel
, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
The sequence entry for this EST has been reverse complimented and
is being submitted in the sense orientation.
Seq primer: M13-21.
FEATURES
source
1. 389
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="120N14XP"
/clone_1ib="Lambda-PRL2"
/note="Vector: lambda zip-lox. Site_1: SalI. Site_2: NotI;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda zip-lox. The cDNA
inserts were directionally cloned with SalI-Not arms using
oligo dT primed cDNA. "
BASE COUNT 88 a 83 c 82 g 130 t 6 others
ORIGIN

```

BASE COUNT	ORIGIN	143 a	114 c	129 g	131 t	21 others
Query Match		20.6%;	Score 235.4;	DB 189;	Length 538;	
Best Local Similarity		90.9%;	Pred. No. 2.5e-53;			
Matches 289;	Conservative	0;	Mismatches 25;	Indels 4;	Gaps 4	
QY 1	aggagaataaaaaaacatgatgatgaggaagaagatcgtggtttgaagcctaagcttaaggt 60					
Db 148	AGGAAAAATAAAAAACATCATCATATGCGGAAAAAGATCTAGAGTTTGACCTTAAGCTTTAGGGT 207					
QY 61	tttcacaaatacaaatccctcttcagatgaatcgaatccaaacttcattatcaaca 120					
Db 208	TTTACAAAATTCACATCTCTTTGAGATGAATCTGAATCTTACTCTTCATTATCAACA 267					
QY 121	atctccagagactcccatggaacaaacattcgatccctacatcagatcttcgaagaag 180					
Db 268	ATCTCAGAGACTCCCATGGGAACAAACATCTCGATCAATCATGATCTTCGCAAGNTAG 327					
QY 181	ac-gtgaacagttttccatcaac-vgttaactgcgaggaagaacacagagtttcgtacc 238					
Db 328	GCGGTGACACAGTTTTCATCAATCAAGGGGTTTAACTGCAGAGNAGCACAGAGAGTTTGTCACC 387					
QY 239	aaacagtagacatctcaagcaccttagtcgaggaagaagatgagagaagaatctcgg 298					
Db 388	AAACAGT-CGGTCTCAAGCACCTTGGGGGAGGAGAGTTGNAG-AGGCAATTTCCGG 445					
QY 299	aaccgcgttgcgtccg 316					
Db 446	ACCCGGGTGNTCCGG 463					
RESULT 11						
B77601	534 bp	DNA	GSS	16-JAN-1998		
LOCUS	T27C23TF	TAMU Arabidopsis thaliana genomic clone T27C23	DNA			
DEFINITION	sequence.					
ACCESSION	B77601					
VERSION	B77601.1	GI:2774240				
KEYWORDS	GSS.					
SOURCE	thale cress.					
ORGANISM	Arabidopsis thaliana					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.					
REFERENCE	1 (bases 1 to 534)					
AUTHORS	Rounsley,S.D., Field,C.E., Baas,S., Linher,K., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter					
	,J.C.					
TITLE	A BAC End Sequence Database for Identifying Minimal Overlaps in					
JOURNAL	Arabidopsis Genomic Sequencing. Update 3					
COMMENT	Unpublished (1997)					
	Other GSSs: T27C23TR					
	Contact: Steve Rounsley					
	Department of Eukaryotic Genomics					
	The Institute for Genomic Research					
	9112 Medical Center Dr., Rockville, MD 20850, USA					
	Tel: 301 838 0200					
	Fax: 301 838 0208					
	Email: rounsley@tigr.org					
	Seq primer: M13-21					
	Class: BAC ends					
	High quality sequence stop: 534.					
	Location/Qualifiers					
	1..534					
FEATURES						
source	/organism="Arabidopsis thaliana"					
	/strain="Columbia"					
	/db_xref="taxon:3702"					

```

/clone="T27C23"
/clone.lib="RAMU"
/sex="hermaphrodite"
/note="vector: BelosaxCt; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"
172 a 102 c 120 g 140 t

```

Query Match	14.7%	Score 167.6;	DB 256;	Length 534;
Best Local Similarity	70.1%	Pred. NO. 7.1e-35;		
Matches 269;	Conservative 0;	Mismatches 109;	Indels 6;	Gaps 13;

477 aaacggaagctagcttctgctaagaagctgaacttgacgycgaagcaagtcgaagtctgg 536
|||||
152 aaacgaagctggccttggcctaagaagctgaagtttaacacgaacagacgaagtcgaagtcgc 211
|||||

537 ttccaaacagagagctagtgaga----ttctctgcttaacccaattgttaaa 592
|||||
212 ttccaaacagagagagctagtgaga----ttctctgcttaacccaattgttaaa 592
|||||

593 ttacattataccccctgtgtagaagtttacc-ttaacaagtcogatttggtttaca 651
||||| ||| | | | | | | | | |
272 GACCAATTTGGCCCCCTAAGAGCAGCATGTAATTTCACAATACACCACATTTTGTGCCTATA 331

652 gaaccâagttaaacgaacggaagttagatctgcgataacttgaacggtgcgtagagaagc 711
Y |||||
332 GCACAAAGTTTAAAGCAGACCGCAAGTGATTCGACGTATTGGAAAAGATGTCTTGAGAAAT 391
|||||

712 taacgaagagaaccgagacttcagaaagagcgtatgtgagcttcgaactcgaagctgt 771
|||||
392 TAACGAAGAGAGATTCGCGCGCTTCAGAAAGA-TCGACGCGCACTTAAGAGCAATTAAAGCGTTT 450
|||||

[illegible]

```

832 agcggtgtggtggtcccatcatcat 855
      | | | | | | | | | | | | | |
511 AACCCGCGCGCGCGCGCGCGCGCG 534

```

RESULT 12

LOCUS	531 bp	mRNA	EST	06-SEP-2000
AV551157				
AV551157	Arabidopsis	thaliana roots	Columbia	Arabidopsis thaliana
cdna clone R2122a06r	5',	mRNA	sequence.	

ACCESSION	AV521157	
VERSION	AV551157.1	GI:8722570
KEYWORDS	EST.	
SOURCE	thale cress.	

ORGANISMISM
Algaecoopisps lialialia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE	AUTHORS	TITLE
1 bases 1 to 531)	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.	A large scale analysis of cDNA in <i>Arabidopsis thaliana</i> : Generation of 12,028 non-redundant expressed sequence tags from normalized cDNA

size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
Contact: Erika Asamien
JOURNAL
MEDLINE
COMMENT

The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: kazusa@kazusa.or.jp, kazusa@kazusa.or.jp

FEATURES	Location/Qualifiers
source	1. 531
	/organism="Arabidopsis thaliana"

```
/db_xref="taxon:3702"
/clone="R2122a06r"
```

BASE COUNT	163 a	115 c	110 g	143 t
ORIGIN	/clone.lib="Arabidopsis thaliana roots Columbia" /tissue.type="roots" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"			

Query Match	14.3%;	Score 162.8;	DB 31;	Length 531;
Best Local Similarity	68.8%;	Pred. No. 1.4e-33;		
Matches 278;	Conservative 0;	Mismatches 102;	Indels 24;	Gaps 3;

QY 9 aaaaacatgataatgtggcacaagagatctaggtttgagcctaaactatgggtttcaca 68
133 aatattcatgatctggcgttaagagatcttggcgtttaactctttacgtcttcgcatttcgacaa 192

QY 69 atcaccatactcttcagatgtagtactgaacacctccatcatcaaacatctccag 128
 ||| ||||| ||| ||| ||| ||| ||| |||
Db 193 AACCC--ATCCTCTCCACTAAATCTTAAACCCACTTCTTCACCATAATGTCATTCTCCAG 249

QY	129	agactcccatggaaccaacatcgcatacctacacaga-----tctt	170
Db	250	ATCTTTTCATGTAACCAAAACCTGTGTTCTTCTCCAGATCAACAAACACAGTTTCTT	309

Oy	171	cgcgaagatagacgctgaacagttlctccalcacaggtlaactgcgaggaacacacaggagt	230
Db	310	AGCAAAATCGACCTGMAACGCTTGCCACCAACAGCTGGATTGTAAGACAGACAGCAAGATT	369

Oy	231	tcttcacccaacagtaagatcctcaagcacatttagcyggaagaagtgtagagagaaga	230
Dh	370	TCTGTCCTCCAAACAGTAGCATCTCGAGCACAGTGTGGAAAAGAGAG--GACTACTATAA	426

OY 291 atctccgaaccgcgcttgcgtccgcgcacgatcacagcagatcaactccgcatcgaggg 350
AGAGAAGGTATCTTCCTCGTGGTGTTGCGTAGACACTTGATCATCACTAGATAGACT 486

Dh
OY

351	tactcacgtgaacccatcagatgaaagaagaacgcggcgcaaac	394
487	TCCCTCACTGGAACTCCTCGATGGAAGAGATTACGGGAAC	530

RESULT 13

LOCUS	569 bp	mRNA	EST	07-SEP-2000
AV5533327		Arabidopsis	thaliana	Arabidopsis thaliana
AV5533327		Arabidopsis	thaliana	Arabidopsis thaliana
CDNA	Clone R260a03R 5'	mRNA sequence.		

```

VERSION      AV553327.1  GI:8724740
KEYWORDS
SOURCE       EST.
             thale cress.

```

CARNIVORAE
 NUDIPEDIS: Entolidae
 Eukariota: Streptophyta; Embryophyta: Tracheophyta;
 Eukariota; Viridiplantae; Streptophyta; eudicots;
 Spermatophyta; Magnoliopsida; core eudicots;
 Rosidae II; Brassicales; Brassicaceae; Arabidopsis.
 Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE	AUTHORS	TITLE
(uses 1 to 303)	Asamitsu, E., Nakamura, Y., Sato, S. and Tabata, S.	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalised and

JOURNAL
MEDLINE
COMMENT
Contact: Erika Asamizu
DNA Res. 7, 175-180 (2000)
Size-Selected cDNA Libraries
20363093

The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: aseai@kazusa.or.jp Or to <http://www.kazusa.or.jp/en/plant/>

FEATURES	Location/Qualifiers
source	1..569
	/organism="Arabidopsis thaliana"

```

/db_xref="taxon:3702"
/clone="R260a03R"
/clone_lib="Arabidopsis thaliana roots Columbia"

```


/tissue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 171 a 122 c 119 g 157 t
ORIGIN

Query Match 13.9%; Score 158.8; DB 31; Length 569;
Best local similarity 69.1%; Pred. No. 1,8e-32;
Matches 271; Conservative 0; Mismatches 97; Indels 24; Gaps 3;

QY 9 aaaaatgatgatggcgaagaagaatctagatttgagcctaagcttaggtttcaca 68
DB 165 AATATCATGATGATGAGGTAAGAGAGATTGGTTAAGTCTTAGCTTGGATTGGCA 224
QY 69 aatcacaatctcttcagatgaatctgaatcctaactcttcattatcaacaatccag 128
DB 225 AACCC--ATCCTTCACGTAATCTTAAACCCACTTCACCAATGTCAATCTCCAG 281
QY 129 agactcccatgagacacaacattcgatcctacatcaga-----tctt 170
DB 282 ATGTTTCATGAGAACCAACCTTGTCTTCTTCAGATCAACAAAGCAACAGTTTCTT 341
QY 171 cgcgaatagacgtgacagtttccatcaacaggttaactgcgaggaagaacacagagtt 230
DB 342 AGGAAATGACGCTGAAACACCTTGCCACACACGATTTGGAAGAGAGACAGAGTT 401
QY 231 tctgtaaccaaacgtatgcattcaagcaccattagcgggaagaagaagtgaagaaga 230
DB 402 TCTCTCCCAACAGTACGATCTGACACGTAAGTGAAGAGAG--GAATTAATGAA 458
QY 291 atctcggaaacgcgttgctccgcgcgacatcacacagacagatcctccgagtcga 350
DB 459 AGGAAAGTACCTCCGCTGCTGCTGCGAGATGACCTTGACATCACTAGATATCT 518
QY 351 tactcaagtgaacctcagatgaagaagaaga 382
DB 519 TCCTCACGTGGAACCTCCGATGAAGAGAGA 550

RESULT 14
T43429 428 bp mRNA EST 07-JAN-1998
LOCUS 6692 Lambda-PR12 Arabidopsis thaliana cDNA clone 120H477, mRNA
DEFINITION sequence.
ACCESSION T43429
VERSION T43429.1 GI:2758294
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 428)
Newman,T., deBuijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL On Jan 7, 1998 this sequence version replaced gi:947835.
MEDLINE Contact: Thomas Newman
COMMENT MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PR1, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcln@clm.cl.msu.edu
Seq primer: T7.
FEATURES
Location/Qualifiers
1..428

/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="120H477"
/clone_1lb="Lambda-PR12"
/note="Vector: Lambda zip-lox; Site_1: Sal; Site_2: Not;
Lambda PR12 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BR1's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 119 a 88 c 100 g 105 t 16 others
ORIGIN

Query Match 12.8%; Score 146.4; DB 189; Length 428;
Best local similarity 64.8%; Pred. No. 4e-29;
Matches 278; Conservative 0; Mismatches 126; Indels 25; Gaps 4;

QY 12 aacatgatgatggcgaagaagaatctagatttgagcctaagcttaggtttcaca 71
DB 1 ATCATATGATGATGGTANAGAGATNGTNTTAAAGTCTTAGCTTGGATTGGCAAAAC 60
QY 72 caaatccctcttagtgaatctgaatcctaactcttcattatcaacaacattccagaga 131
DB 61 C--ATCCTTCACGTAATCTTAAACCCACTTCTACCAATGTCAATCTCCGATG 117
QY 132 ctcccatggaaccaacattcgatcctacatcaga-----tcttcgc 173
DB 118 TTTCATGAGAACCAACCTTGTCTTCTTCAGATCAACAAAGCAACAGTTTCTTAG 177
QY 174 aagatagacgtgacagtttccatcaaacggttaactgcgaggaagaacacagagttcg 233
DB 178 AAATGACGTGAAACAGCTTGCCACACAGGTGATTTGGAAGAGAGACAGAGTTTGC 237
QY 234 tcacaaacagatcgatctcaaacgaccatttagcggaagaagaagtgaagaagaatc 293
DB 238 TCTCCAAACAGTACGATCTCGNCACAGTGAAGAGGNGAGACTCTGAAAGAA 297
QY 294 tccggaacgcgttgctccgcgcgacatcacagacagatcactccgagtcgaagttac 353
DB 298 GGTACCTCGGNTGCTGTTGCGAGATGACCTTGAC--ATGACTCTGATGATGTTCC 354
QY 354 tcaagtgaacctcagatga-agaagaagaacggtggcgaaacgtcgagaagaagctcag 412
DB 355 TCACGTGGGACCTNCNTGAGAGAGCTTACGAGGTTGNCNTTANGAAGGAAGTTAG 414
QY 413 gttatcaaa 421
DB 415 NNTTCCCAA 423

RESULT 15
BG441022 860 bp mRNA EST 15-MAR-2001
LOCUS BG441022
DEFINITION GA_Ea0011F23f Gossypium arboreum 7-10 dpa fiber library Gossypium
ARBOREUM cDNA clone GA_Ea0011F23f, mRNA sequence.
ACCESSION BG441022
VERSION BG441022.1 GI:13350674
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 860)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution

JOURNAL
COMMENT

of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293

Email: rwing@clemson.edu
Seq primer: TAATACGCTCACTATAGG
High quality sequence stop: 746.

FEATURES

source

location/Qualifiers
1. 860

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0011F23f"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"

BASE COUNT 257 a 190 c 222 g 186 t 5 others
ORIGIN

Query Match

12.3% Score 140.8; DB 153; Length 860;

Best Local Similarity 66.0%; Pred. No. 1.6e-27;
Matches 268; Conservative 0; Mismatches 102; Indels 36; Gaps 3;

```

OY 152 cgatcctacatcagatcttcgaaagatagacgtgaacagtttccatcaacggttaactg 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 CGAGACTAGATCATTTCTTAGAGAGATCGACGTGACAGATTGCCATCTACGTCGATTG 66

OY 212 cgaagaagacacagagatttgcaccacaagatagatctcaagcaccattagcggaa 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 CGAAGAAAGAGCTGAGTTTCATCTCCGACACGACGATATCGAG---TGTAGTGGGAA 123

OY 272 gagaagtgaagaagaatctcggaaacccgcgttgcgcgcgacgacacga 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 AAGGAGCGAAGGAGGACACTATATGA-----GATGA 156

OY 332 gatcactcggatcgaaggtactcaacgttgaacctcagatgaagaagaagcggggcga 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 GCTTGATATTGAAGAGCTGTCTCGTGGCATCATGATGAGGAAGAT-----GGTGA 210

OY 392 aacgtcgaagaagaagctcaggtatatacaaaagatcagtcgtcttctcgaagaagactt 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 CGTTTCGAGAAAAAACTTAAGCTTTTCGAAGGATCAGTCTGTATTCTTGAAGAAAGCTT 270

OY 452 caaagaacacacactcctcaacaaagaagaagcttagcttggctaaagaagctgaact 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 TAAAGAACACACACTCTGAACCCAAACAAACAAAGATGCTTGGCTAAGCAGCTGGGATT 330

OY 512 gacggaagaagaagtggaagtgtgttccaaagaagaagactag 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 GCGACCCGAGCAACTTGAAGTTGTTCCAAACAGAAAGGCGGAG 376
  
```

Search completed: June 19, 2001, 17:00:51
Job time: 22534 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:28 ; Search time 7917.7 Seconds
(without alignments)
1212.991 Million cell updates/sec

Title: US-09-394-519-43

Perfect score: 1016
Sequence: 1 tgggaagccacaataacc...aaaaaaaaaaaaaaaa 1016

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.
1 (bases 1 to 587)
Newman,T., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers

FEATURES
source
1..587
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="193M157"
/clone_1lb="Lambda-PRU2"
/note="Vector: lambda Zip-lox; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT
ORIGIN
165 a 106 c 135 g 167 t 14 others

Query Match 45.3%; Score 460.6; DB 158; Length 587;
Best Local Similarity 94.2%; Pred. No. 1.3e-83;
Matches 505; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

QY 2 gggaagccacaataacccctatctcctggcctttttaaaagttaagaataatccg 61
DB 3 GGGNNCCACAAATMACCCCTATTCCTCGCCCTTTTAAAAAGTTTAAAGATAATCCG 62
QY 62 ataaatctctatataatatttcttggtcgaagaggttcgccaagaggttga 121
DB 63 ATTAATCTTTTATTTATTTTNTTGGTCCATGAGAGGTTCCCAAGGGTTGAG 122
QY 122 gaaagtgatgactgctgaaagaagatagctctctgaggtatgataagatag 181
DB 123 GAAAGTGATGACGTGGAAGAAGATAGTCTTGAGGCTATGTATGATAGTATGG 182
QY 182 agaaagcaatgcatcaagttccttgagagtggtggtcctaactcgatcgagaaggttg 241
DB 183 AGAAGCAATGCGATCAAGTTCCTTTGAGAGCTGGCTTAATCGATGCAAGAAAGTTG 242
QY 242 tagaactaagatgctgaactattgagccaagtatcaagaagaggaagactgacatg 301
DB 243 TAGACTAAGATGTTTAACTATTGGAAGCCAGTATCAAGAGGAGAGACTTATGCATGA 302
QY 302 tgaagttagctctctctccttcataaagctctcagaagaata-9gtgctcctgattg 360
DB 303 TGAAGTTGATCTTCTTCGNCCTCATTAAGCTTGAAGAAATAGGTTGGTCTTATG 362
QY 361 ctggtctgactgctggtcgagccgctaagatgtaaaata-ctggaaccacatcg 419

DB 363 CTGTCGATTGCTGTCGCGACCGCTAATGATCAAAAATTACGTGAACACCATCTG 422
QY 420 agtaaaaaacatgagctctcggttgtaagcttaaatgaagaagaacatttttc 479
DB 423 AGTAATAACATGAGNGCTTGCTGTGTAAGGCTTAATATAAAGGAAACCATATTTC 482
QY 480 cctctcacaca-ccgctcaaaaacatcggtttttaagctcgacatcgatcct 534
DB 483 CCTCCTNCAACANCCGCGNCAAAAANCGGTGTTTAAAGCTTGNCCTGATGCCCT 538

RESULT 3.
LOCUS T22223
DEFINITION 4231 Lambda-PRU2 Arabidopsis thaliana cDNA clone 99P97, mRNA
sequence.
ACCESSION T22223
VERSION T22223.1 GI:2596878
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.
1 (bases 1 to 382)
Newman,T., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On Nov 6, 1997 this sequence version replaced gi:932253.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers

FEATURES
source
1..382
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="99P97"
/clone_1lb="Lambda-PRU2"
/note="Vector: lambda Zip-lox; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT
ORIGIN
104 a 64 c 89 g 110 t 15 others

Query Match 23.3%; Score 236.6; DB 188; Length 382;
Best Local Similarity 82.7%; Pred. No. 3.2e-38;
Matches 286; Conservative 0; Mismatches 57; Indels 3; Gaps 2;

QY 71 tttaataaattttcttggtcatalggaaggttcgccaagaggttgagaaggtgc 130
DB 23 TATTTAACGATCATCTTGTTCATGAGAGGTTGCTCAAGGGCTCGCAAAAGTGC 82
QY 131 atgactctgaaagaagatagctctcttgaggtatgtatgataagatgagaagca 190

LOCUS BE054276 686 bp mRNA EST 07-MAR-2001
 DEFINITION GA_Ea002018f Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Ea002018f, mRNA sequence.
 ACCESSION BE054276
 VERSION BE054276.2 GI:13244153
 KEYWORDS EST.
 SOURCE Gossypium arboreum.
 ORGANISM Gossypium arboreum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
 1 (bases 1 to 686)
 Wing, R.A., Firsich, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
 D., Wood, T.C., Leslie, A. and Wilkins, T.A.
 An integrated analysis of the genetics, development, and evolution
 of the cotton fiber
 Unpublished (2000)
 On Jun 8, 2000 this sequence version replaced gi:8381332.
 JOURNAL Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall
 Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATAGACTACTATAGG
 High quality sequence stop: 559.
 FEATURES
 source
 1..686
 /organism="Gossypium arboreum"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_Ea002018f"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /tissue_type="Fibers isolated from bolls harvested 7-10
 dpa"
 /lab_host="E. coli"
 /note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 228 a 142 c 146 g 170 t
 ORIGIN
 Query Match 13.7%; Score 139.6; DB 162; Length 686;
 Best Local Similarity 65.3%; Pred. No. 1.5e-18;
 Matches 205; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 Oy 113 aggggtgaaggaaggtgcatgactgctgaagaagatagctctcttgaggtatgtatga 172
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 155 AGGGTGAAGAAAGGTCCATGAGACCCAGAAAGATCAAAAGCTTACCTTACATTGA 214
 Oy 173 taagtgtggaaggaagaaagcatcaagtctctcttgaggtggtggttaataatgatacag 232
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 215 ACAACATGGCCATGAGAGCGTGGCTTGCCTTTAAAGCTGGGCTTCAAAAGATGTGG 274
 Oy 233 aaagagttgtagactaagatggttgaactattgaagccaagatatacaagaaggaagact 292
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 275 AAAGAGTTGAGAGATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 334
 Oy 293 tagcaatgtagaagttgatactctctcttgagcctcaataagcttcaagaataagtggtc 352
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 335 CAGTTTACAAAGAAAGACAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 394
 Oy 353 ctgattgctggtcgtgactgctggtgacgcgttaagtgatgataaatacttgagaaacac 412
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 395 TGGCATTAGCTACTTATTTGCCGAAAGACAGACAAATGAATGAATGAATGAATGAATGAATGAAT 454
 Oy 413 ccacttgagtaaaa 426
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 455 ACATCTTAAGAAAA 468
 RESULT 12

AW255172
 LOCUS AW255172 621 bp mRNA EST 23-AUG-2000
 DEFINITION M160 peppermint glandular trichome Mentha x piperita cDNA, mRNA
 sequence.
 ACCESSION AW255172
 VERSION AW255172.1 GI:7244424
 KEYWORDS EST.
 SOURCE peppermint.
 ORGANISM Mentha x piperita
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Lamiales; Lamiales; Mentha.
 1 (bases 1 to 621)
 Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D. and
 Crobeau, R.
 Probing essential oil biosynthesis and secretion by functional
 evaluation of expressed sequence tags from mint glandular trichomes
 Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2934-2939 (2000)
 JOURNAL 2018392
 MEDLINE
 COMMENT Contact: Lange, B.M.
 Institute of Biological Chemistry/Washington State University
 Pullman, WA
 Email: lange-m@mail.wsu.edu.
 FEATURES
 source
 1..621
 /organism="Mentha x piperita"
 /cultivar="Black Mitcham"
 /db_xref="taxon:34256"
 /clone_lib="peppermint glandular trichome"
 /tissue_type="pelate glandular trichomes"
 /cell_type="secretory"
 /note="Vector: lambda ZAPit"
 BASE COUNT 204 a 130 c 163 g 124 t
 ORIGIN
 Query Match 13.5%; Score 137.2; DB 113; Length 621;
 Best Local Similarity 63.8%; Pred. No. 4.5e-18;
 Matches 208; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
 Oy 113 aggggtgaaggaaggtgcatgactgctgaagaagatagctctcttgaggtatgtatga 172
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 60 AGGGCTGAAGAGAGGAGATGAGACTGCTGAAGAGACAAACTAAGAAATTAATTC 119
 Oy 173 taagtgtggaaggaagcaatggtcctgaagctctcttgaggtggtggttaataatgatacag 232
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 120 GGAAGATGGTGAAGCGTGGTGGAGATCATTTGCCCAAGAAATGCAAGTTTACTTATGATGTGG 179
 Oy 233 aaagagttgtagactaagatggttgaactattgaagccaagatatacaagaaggaagact 292
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 180 AAAGAGTTGAGAGATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 239
 Oy 293 tagcaatgtagaagttgatactctctcttgagcctcaataagcttcaagaataagtggtc 352
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 240 TTCTTCGAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 299
 Oy 353 ctgattgctggtcgtgactgctggtgacgcgttaagtgatgataaatacttgagaaacac 412
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 300 CTTGATGCCGCGACTTGGCGGTAGACAGACAAATGAATGAATGAATGAATGAATGAATGAATGAAT 359
 Oy 413 ccacttgagtaaaaacatgagttct 438
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 360 CCATTGAGCAGAAATTTTCACGGTT 385
 RESULT 13
 BG134230
 LOCUS BG134230 447 bp mRNA EST 31-JAN-2001
 DEFINITION EST467122 tomato crown gall Lycopersicon esculentum cDNA clone
 ctoel5D4 5' sequence, mRNA sequence.
 ACCESSION BG134230
 VERSION BG134230.1 GI:12634418
 KEYWORDS EST.

SOURCE	tomato.
ORGANISM	Lycopersicon esculentum
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids I; Solanales; Solanaceae; Solanum;
AUTHORS	1 (bases 1 to 447) <i>I Lycopersicon.</i> van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Hansen,C., Ronning, .C. and Tankley,S. Generation of ESTs from tomato crown gall tissue Contact: CUGI Unpublished (2001)
TITLE	Unpublished (2001)
JOURNAL	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html.
COMMENT	
FEATURES	Location/Qualifiers 1..447 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="CTOE1SD4" /clone_lib="tomato crown gall" /tissue_type="crown gall" /dev_stage="crown galls from full-grown plants (8 wks old)")lab_host="SOLR" /note="vector: pBluescript SK(-); Site.1: EcoRI; Site.2: XhoI; Four wk old greenhouse plants were stab inoculated on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr Cornell U.). Galls were allowed to develop for another 4 wks, when gall tissue was frozen in liquid nitrogen."
BASE COUNT	163 a 69 c 98 g 117 t
ORIGIN	
Query Match	. 13.5%; Score 137; DB 174; Length 447;
Best Local Similarity	63.7%; Pred.No. 4.9e-18;
Matches 225; Conservative	0; Mismatches 125; Indels 3; Gaps 1
OY	103 gtctgccaaaggttgaggaaaagtgcataggacgctgaagaatattcttcctaagc 162
Dd	74 GTTCTAAGGAAGGTTAAAGAAGGTTCATGGACATAAAAAGAAGATTGCCTCTAACG 133
OY	163 tatgatcataagtatgagaagcaaa--atgcatcaaagttcccttgagaagctggc 219
Dd	134 AATATATTAAAGCAATATGGTGGAAGACTCGTAGTGAGATCTCTCTTAAAAAGCACGGCC 193
OY	220 taatatcgatgcagaagaagtgttaactaagaatggttgtaacctttgaagccaagratca 279
Dd	194 TACTAAGATGTGCGACAAGACTTGTACATTAAAGATGGGTGAATTTATCTTAAAGCTTGACACTCA 253
OY	280 agagaaggaagacttagcaaatgaatgaattgattcttccttcctgccttcataagctctag 339
Dd	254 ATAGAGGCGACCTTTACACCAGCAAGAGAGACAGACTTATTGTTGCACCTTTTACCTTTTGG 313
OY	340 gaataagtggtgaccttgatttgttcgtatgcattgcctggctgaccgccaatgltgtaaaaa 399
Dd	314 GTAGTCGTTGGTCTCTCATTTGCCGGAGAGATACCTGGTGAACAGCAATGACATCAAAGA 373
OY	400 attactggaaccaccatctgaagtaaaaaaaaaacaagatgcttcgctggtttaagctc 452
Dd	374 ATTATTGGAACACGAGTCTCTCCAAGAAACCTAAAGCTCAAGAAATGAGCCCT 426
RESULT 14	
LOCUS	AM278294 355 bp mRNA EST 04-JAN-2000
DEFINITION	sfr2c03.yl Gm-cl009 glycine max CDNA clone GENOME SYSTEMS CLONE ID
TRANSCRIPTIONAL FACTOR MBFL	Gm-cl009-7621 5' similar to FR:Q40920 Q40920 MYB-LIKE
VERSION	AW278294.1 GI:6666835

KEYWORDS	EST. soybean.
SOURCE	Glycine max
ORGANISM	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucotsids I: Fabales: Fabaceae: Papilionoideae: Glycine
REFERENCE	1 (bases 1 to 355)
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Corryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,T., Allen,T., Bowers,R., Riltter,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Trace considered overall poor quality Seq primer: -40RP from Gibco High quality sequence stop: 1.
FEATURES	Location/Qualifiers 1..355 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl009-2621" /clone_id="Gm-cl009" /lab_host="X10-Gold" /note="Vector: pluescript II XR, site_1: EcoRI, site_2: XhoI; The mRNA was isolated from entire roots of 2-month-old 'Williams' plants that were greenhouse grown in 5-gallon pots. To suppress nodulation, Black Gold All-Purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (P205), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammonia N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used (GAACAGAGAGAGAGAGAGACATCTCAG(T)-18). After second-strand synthesis is, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSep 400 Spin column from Pharmacia. The column eluent was then ligated into Stratagene's pluescript II XR predigested vector (pluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."
BASE COUNT	130 a 69 c 79 g 77 t
ORIGIN	
Query Match	13.3%; Score 135.2; DB 114; Length 355;

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** Small print

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)
